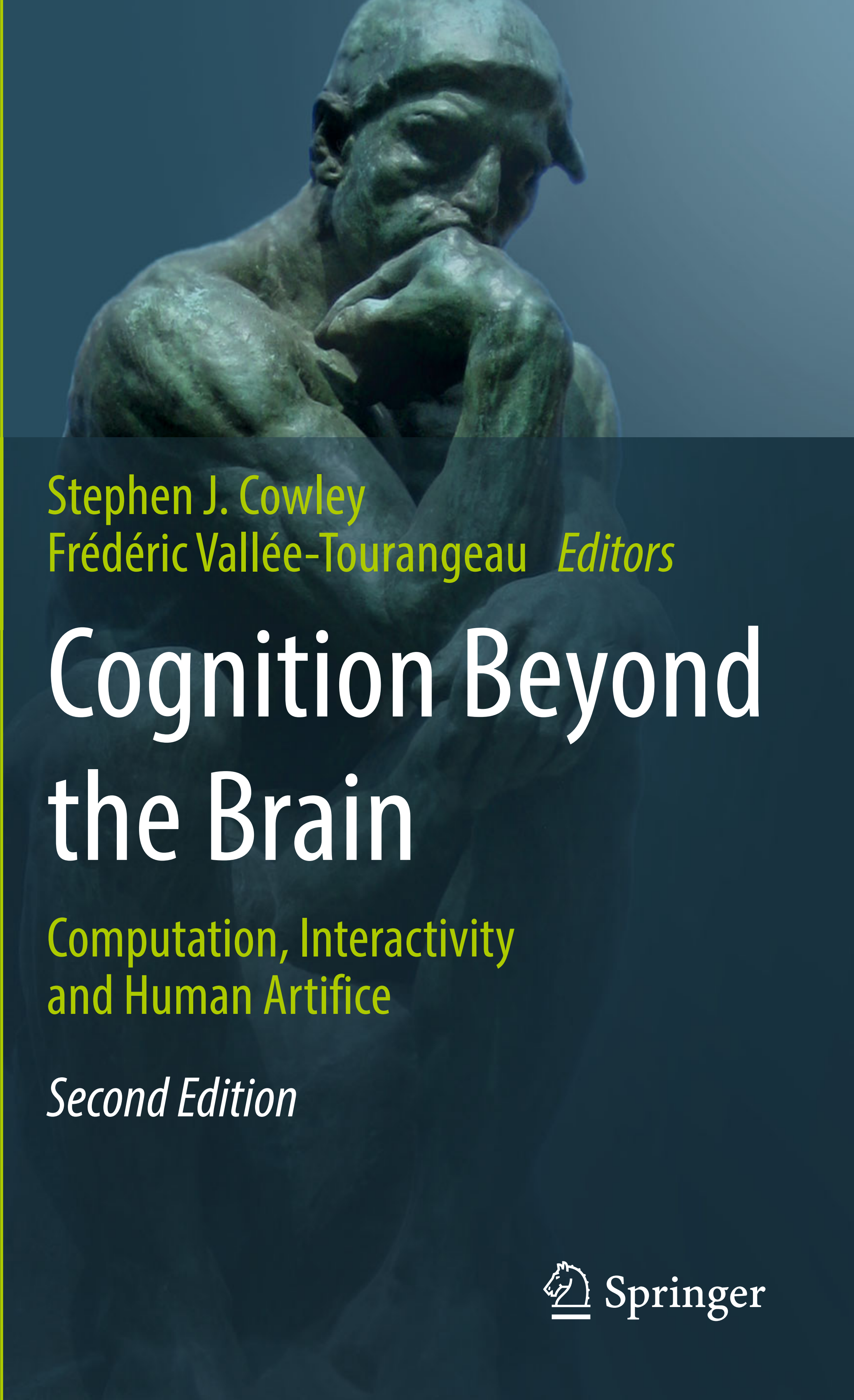


Stephen J. Cowley
Frédéric Vallée-Tourangeau *Editors*

Cognition Beyond the Brain

Computation, Interactivity and Human Artifice

Second Edition

Springer

Cognition Beyond the Brain

Stephen J. Cowley · Frédéric Vallée-Tourangeau
Editors

Cognition Beyond the Brain

Computation, Interactivity and Human Artifice

Second Edition

Springer

Editors

Stephen J. Cowley
Department of Language
and Communication
University of Southern Denmark
Slagelse
Denmark

Frédéric Vallée-Tourangeau
Department of Psychology
Kingston University
Kingston upon Thames
UK

ISBN 978-3-319-49114-1 ISBN 978-3-319-49115-8 (eBook)
DOI 10.1007/978-3-319-49115-8

Library of Congress Control Number: 2016961336

1st edition: © Springer-Verlag London 2013
2nd edition: © Springer International Publishing AG 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Contributors

Billy Adamsen Zealand Institute of Business and Technology, Slagelse, Denmark

Alexander Aston School of Archeology, Oxford University, Oxford, UK

Chris Baber Electronic, Electrical and Computer Engineering, The University of Birmingham, Birmingham, England

Linden J. Ball University of Central Lancashire, Preston, UK

Jonathan Ben-Naim CNRS and Université de Toulouse, Toulouse, France

Jean-François Bonnefon Toulouse School of Economics, Center for Research in Management, Institute for Advanced Study in Toulouse, Centre National de la Recherche Scientifique, Toulouse, France

Robert W. Clowes Faculdade de Ciências Sociais e Humanas, Instituto de Filosofia da Nova, Universidade Nova de Lisboa, Lisbon, Portugal

Stephen J. Cowley Centre for Human Interactivity and the COMAC Cluster, University of Southern Denmark, Slagelse, Denmark

Richard Heersmink Department of Philosophy, Macquarie University, Sydney, Australia

Andreas Herzig CNRS and Université de Toulouse, Toulouse, France

David Kirsh Department of Cognitive Science, University of California, San Diego, La Jolla, USA

Sylvie Leblois Université de Toulouse, Toulouse, France

Josef Lhotský Department of Philosophy and History of Science, Faculty of Science, Charles University Prague, Prague 2, Czech Republic

Damien Litchfield Edge Hill University, Ormskirk, UK

Emiliano Lorini Université de Toulouse, Toulouse, France

Jens Koed Madsen School of Geography and the Environment, University of Oxford, Oxford, England

Lambros Malafouris Keble College and the Institute of Archeology, University of Oxford, Oxford, UK

Anton Markoš Department of Philosophy and History of Science, Faculty of Science, Charles University Prague, Prague 2, Czech Republic

Martin Neumann Jacobs University Bremen, Bremen, Germany

Mark Perry Department of Computer Science, Brunel University, Uxbridge, UK

Kåre Stokholm Poulsgaard Institute for Science, Innovation and Society, University of Oxford, Oxford, UK

Davide Secchi COMAC Research Cluster, Centre for Human Interactivity (CHI), Department of Language and Communication, University of Southern Denmark, Slagelse, Denmark

Paul Smart Electronics and Computer Science, University of Southampton, Southampton, UK

Sune Vork Steffensen Department of Language and Communication, University of Southern Denmark, Odense, Denmark

Jana Švorcová Department of Philosophy and History of Science, Faculty of Science, Charles University Prague, Prague 2, Czech Republic

Frédéric Vallée-Tourangeau Department of Psychology, School of Social and Behavioural Sciences, Kingston University, Kingston, UK

Gaëlle Vallée-Tourangeau Department of Management, Kingston Business School, Kingston University, Kingston, UK

Chapter 1
Thinking, Values and Meaning in Changing Cognitive Ecologies

Stephen J. Cowley and Frédéric Vallée-Tourangeau

Abstract In defence of the pluralism of *Cognition Beyond the Brain,* we propose a systemic view of cognition. We argue that this opens up the diversity of thinking by allowing persons to act as living subjects who also make routine use of collectively organised cognitive ecosystems (Hutchins 2014). To allow for the duality of human agency, one needs a methodological principle that links subjective concerns with collective control. The principle of cognitive separability (PCS) posits that people can dissociate extended causal processes from subject-centred actions. The systemic view offers three main claims about human cognition. First, language, problem solving, human-computer interaction, and many elaborate cognitive skills are based in sense-saturated coordination or *interactivity*. Second, interactivity prompts embrained bodies to develop the cognitive powers used in human performance. Third, human agents entwine brains, bodies and world as they use the PCS to manage multi-scalar dynamics. In this second edition of *Cognition beyond the brain*, we defend the principles by presenting the state of the art. The papers address, among other things, use of the internet, the fluid nature of self, the history of the Paris Commune, and computer aided design. In line with cognitive separability, the studies vary in the granularity used to pursue how bodies and artifacts shape thinking. Whereas some contributors invoke (putative) mental content, others take the radical view that thinking depends on non-localisable pattern. Yet, in all cases, experience of materiality is meshed with activity that involves the world beyond the body.

S.J. Cowley (✉)
Centre for Human Interactivity and the COMAC Cluster,
University of Southern Denmark, Slagelse, Denmark
e-mail: cowley@sdu.dk

F. Vallée-Tourangeau
Department of Psychology, School of Social and Behavioural Sciences,
Kingston University, Kingston, England
e-mail: f.vallee-tourangeau@kingston.ac.uk

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_1

Thinking in Action

Folk tradition traces thinking to an individual who struggles alone with abstract problems (curiously, the penseur is typically a man). The thinker's mark is the furrowed brow shown in Rodin's famous statue. Yet this is thinking of an unusual kind that evokes unusual types of task that are, it seems, limited to some cultural settings. Once the question is raised, abundant evidence shows that thinking uses what lies *beyond* the brain. Above all, it depends on bodies that connect history with agent-environment interactions. As in sport, cooking and language activity, human thinking shows ecological uniqueness. Whether using prehistorical artifacts (Malafouris 2013), making economic choices (Ross 2007) or using written signs (Kravchenko 2009), material entities grant a collectivity some control over how people think and act. The individual-collective nature of human agents shapes the values and meanings that, for James Gibson, are central to psychology. The axiological dimension of life prompts people to link expertise, subjectivity, action and contingently driven, situated use of human artifice. Challenging both folk views and dictionary definitions, we thus deny that thinking is organism-centred or that living brains do more than *enable* human forms of life.

Thirty years ago, no-one looked at how thinking uses resources beyond the brain. Since then a torrent of work has shown how skills and dispositions spread in time as persons use resources on both sides of the skull. Cognition is embodied and situated—as is accepted by even a conservative wing of cognitive science (see, Rupert 2010). In the second edition of *Cognition beyond the brain*, human flexibility is thus traced to simplex principles found in all DNA-based biosystems. Yet, while even viruses link structures across time scales, human subjects use history as they engage with things, events and situations. Following Hutchins (2014), they act in cognitive ecologies whose outputs presuppose social, institutional and technological organisation. To clarify the powers of *human cognitive agents* (Giere 2004), we propose a principle of cognitive separability. While cognitive ecosystems shape socialisation and much activity, persons interpret outputs as they shift between acting as automata and performing deliberately. Human cognitive agents use the resources of reason (e.g., tools, customs, language etc.) to coordinate while using recall, displays of expertise, and individual gleanings of culturally sanctioned knowledge.

The Principle of Cognitive Separability

Cognition beyond the brain traces intelligent activity to bodies that connect persons with nature as metabolism enables them to use objects/dynamics in situations that depend on things, persons and events. To address this variety, Cowley and Vallée-Tourangeau (2013) propose a principle of cognitive separability (PCS). As a methodological principle, it allows human activity—and language—to be traced to dual control. Whether together or alone, people link materiality (a flow of circumstances and objects that constitute contexts) with ways of orienting to persons, things and events (situated social activity). Dual control emerges as infants connect automatic activity with more deliberate acts and, using inhibition, mesh different modes of action. In developmental time, children link sensorimotor skills with historically and culturally mediated use of social norms, artifacts, language, practices etc.

Adopting the principle of cognitive separability has two main advantages. First, in contrast to functionalism or enactivism, it allows for pluralism and debate. This arises from embedding subjective aspects of human agency in how persons use extended systems. According to the PCS, persons can perform like nonhuman agents—as Latour's actants—and, if they choose, draw on culturally sanctioned forms of "enactive signification" (Malafouris 2013). In linking extended and person-centred powers, we connect Järvilehto's (e.g. 1998, 2009) systemic psychology with Giere's (2004) perspectival view of cognition. Human ways of acting/perceiving are 'human centred but not human bound' (Giere 2011). People link products of assimilation (i.e., statistically measurable learning) to learning from behavioural results (Järvilehto 2009). Though lacking space for detailed discussion, cognitive separability seems to use a developmental transition from what Winnicott (1969) calls reliance on 'object-relating' to coming to 'act with objects'. Since these are situated, we call them *things* that, potentially, can shape events as a person's subjective concerns are linked with the doings and beliefs of others. As Winnicott says, "To use an object the subject must have developed a *capacity* to use objects" (1969:712). This capacity is not inborn and, for Winnicott, "nor can its development in an individual be taken for granted. The development of a capacity to use an object is another example of the maturational process as something that depends on a facilitating environment" (Winnicott 1969:712).

A dramatic image of dual control is provided by the operator of a military drone. In such a case, the PCS extends to how missile systems were conceived in terms of cybernetics. First, as in cybernetics, a self-guided drone can hit a target and, as in second-order cybernetics, the drone operator will use an engineering model to manage missile performance. To bring about destruction, no more is required. Yet, in fact, control pertains to an expert who has subjective concerns. While using the output of an extended system (a drone), as a subject, the operator can judge if what is visible is likely to be a wedding or a military target. Thanks to the principle of cognitive separability, a human subject links control with responsibility (or not). While the case may seem extreme, the pluralism of *Beyond the Brain* demands a framework where people link instrumental and more responsible ways of acting. On a weak or situated view of cognition, individual powers can be minimised by positing say, that the internet favours user choice-making (e.g. Smart et al., this vol.). By contrast, a strong or non-localist view grants importance to human subjectivity. Cultural events can prompt a population to become historical actors (e.g., Aston this vol.) or, at other times, a sketch can encourage individuals to rework geometrical knowledge (Kirsh this vol.). Our pluralism fits folk psychology in that use of the internet, doing maths, or raising a revolutionary standard, are all said to depend on *thinking*.

From There to Here

A sea-change took place in cognitive science in the 1990s. Rather than tie cognition to bounded rationality and information processing, attention turned to how the world beyond the brain contributes to cognition. The change comes, in part, from how work in robotics and neuroscience is bound to view action, perception and attention as entwined with language and thinking (not as purely causal). Further, once brains cease to be imagined as symbol processors, symbols disappear as does the search for co-variance between 'faculties' and models of memory, grammar or vision. Rather, one needs a history of cognitive engagement that extends systems biology. Indeed, Rupert's 'active internalism' shows that embodied and situated views of cognition can be reconciled with old fashioned functionalist ideas. In looking beyond the brain, we avoid philosophical issues of intentionality by focusing on empirical work. As the volume testifies, human cognitive agents make much use of interactivity. It is plain that thinking-in-action uses interactivity *qua* sense saturated body-world coordination. Cognitive events arise in space-time as people use chronotopic resources that enable bodies to create, manage—and discuss—values.

Although embodied and situated cognition is not computation, it can make use of von Neumann machines. Thus while philosophers ask whether mind is 'extended' or if an enactivist paradigm will reshape cognitive science, our approach has pragmatic roots. While human cognition is embodied, embedded, enacted and ecological, products like languages and computers *also* manifestly influence thinking. On the systemic approach, human meanings and values draw heavily on cultural products and the world of cultural ecosystems (Hutchins 2014). In using data from the Hubble, for example, complex decision making arises because people use circumstances, make predictions, act and, at times, use counterfactuals. Our view echoes Giere's (2004) comparison between human cognition and farming:

> A distributed cognitive system is a system that produces cognitive outputs, just as an agricultural system yields agricultural products. The operation of a cognitive system is a cognitive process. There is no difficulty in thinking that the whole system, no matter how large, is involved in the process. But there is also no need to endow the whole system with other attributes of human cognitive agents. (Giere 2004:771).

To varying degrees, extended systems are managed by human cognitive agents. While cognitive ecosystems like cockpits and designated crime scenes have a crucial role, much depends on living persons too. In pursuing the heterogeneity of thinking, we connect what is routine and habit-based with subjective judgements. Much depends on what Jaynes (2000) calls excerptions—breaking with an ongoing flow of action to recall the past, imagine the future and, at times, change what happens. Given such skills, people can organise to change the world. Our interest in human cognitive agency draws, in part, on our histories—one of us began as a linguist and the other as a cognitive psychologist; in part, it reflects on a worry that 'relevance' masks important values.[1] But, the main reason for the principle of cognitive separability is that dual control is central to what is human. Though people can treat tools, language and customs as 'autonomous', collectivities rely on techniques and number systems to shape a common world (e.g., one that uses pottery and trade). With population level thinking, shared judgements allow, for example, styles, arithmetic, object names, or 'fairness'. Further, with human language, judgements can be supported by chronicles, narratives and rationalisations. Population level 'thinking' draws on co-acting that is—and is modified by—individual doings. People and communities pursue and evaluate cooperation. As with language or economics, 'thinking' applies to both populations and individuals: just as with 'languages' or the 'macro-economy', it bears on value and experience. In cognitive science, human cognitive agents use the logic by drawing on interactivity.

The systemic view can be traced to a 2009 symposium on distributed thinking at the University of Hertfordshire. Rather than focus on embodiment or situatedness, we asked how cognitive tasks drew on object use. Several of the papers presented there were later published in a Special Issue of *AI and Society* on how cognition is affected by co-acting assemblages (Cowley and Vallée-Tourangeau 2010). In this state of the art, we republish Baber's and Perry's papers from the 2009 symposium: these use distributed cognition to pursue activity a construction site and a (simulated) crime scene. However, we were soon convinced that appeal to co-acting assemblages marginalises persons who, as Peter Jones argues, are the only *active* part of cognitive systems (Jones 2013). Indeed, appeal to co-acting assemblages allows even computer mediated trust (Ben Naim this vol.) to count as sociocognitive. Accordingly, a 2011 symposium at Kingston University turned to what culturally distributed cognition implies for human agency. At this event, we realised that, because human cognition is encultured, it is inseparable from what Kirsh (1997) had called *interactivity*. The claim is supported by Ball and Litchfield, Steffensen and Kirsh (republished here) and used to propose a research programme by Vallée-Tourangeau and Vallée-Tourangeau (this vol.). Alongside interactivity, it is crucial that, in Perry's (this vol.) terms, cognition can be 'demand led' or anticipatory. In today's terms, this is due to the PCS: People use interactivity to link circumstances with cognitive ecosystems in anticipating what will happen. As Giere (2011) stresses, reaching a cognitive outcome—finishing a project—is quite different from modelling its causal basis. There is, he suggests, a 'fundamental asymmetry' between causal loops and their use which enables persons to complete tasks with partial or incomplete understanding. In the case of the Hubble telescope, for example, its causal relays use input and output that is separate from human

---

[1] 'Relevance' tends to be used when mainstream values and discourse are used to justify the parochial concerns of individuals, organizations and political bodies.

bodies.[2] Nonetheless, not only does its design depend on interactivity and its products but this flow of action-perception is also needed to ask questions of the universe. Knowledge results from meshing this with a history of engaging with cognitive ecologies (e.g., families, schools, and science). The situated and the non-localisable (or non-local) dominate thinking. The claim applies to all domains of meaning or value. For example, even being (say) a Tottenham or Arsenal supporter will change experience of many things, events and situations.

The asymmetry of doing and knowing links the principle of cognitive separability to what Malafouris calls "enacted signification". Objects become things linked to networks of routines, procedures and observations that grant them value. By separating observers from the world, language-using beings act *without* any need to understand the causal loops that sustain many actions. We construct a world where it is 'normal' to drive cars, use herbal remedies, consult ancestors or destroy forests. Linguistic powers favour the sense of familiarity used to master techniques and technologies: people draw on talk to probe for solutions, identify problems and explore the adjacent possible. Further, many manifestly fail to separate causal process from interactivity. As a result, it is (mistakenly) posited that brains use in-built logic or act as grammatical engines. Further, Giere's fundamental asymmetry allows extended processes to generate outcomes based on historically derived potential—affordances for possible actions. Even digits (4, 2) may be put together in a question (2 + 4 = ?), used to number a house (24), as part of a code (let 4, 2 = dog) or, as in the *Hitchhiker's Guide to the Galaxy*, to index the meaning of life (42).[3] Cultural products can influence events by promoting observers to link substantive information with various functions. While using physical invariants or statistics, persons also draw on experience as interactivity connects them to a multi-scalar history. From talk to human star-gazing, living systems use biocultural information that grants them modes of action under dual control.

To pursue how people use cultural history, a symposium at Goldsmiths College (in 2013) linked cognitive processes with temporal distribution and, specifically, how slow cultural processes reshape activity in faster scales. The idea appears *in nuce* as human development leads to increasing use of the resources of reason (Neumann and Cowley, this vol.) or indeed how the chromatin system integrates different rates of change (Markos et al., this vol). Accordingly, we pursue the multi-scalar nature of human subjectivity (Madsen, this vol.), organisations (Secchi and Adamsen, this vol.), internet use (Smart et al., this vol.) and enactment of the Paris Commune (Aston this vol.). In each case just as cultural products shape emergent processes, a history of slow scale activity (e.g., making money, Franco-Prussian rivalry) retunes the neural networks used in cognitive action-perception skills. Not only do 'immergent' processes change persons but they make

---

[2] "Nevertheless, there remains a fundamental asymmetry in my view. I want to call a system cognitive because it produces cognitive outputs, but refuse to call it knowledgeable because it produces knowledge. For me, the latter makes as little sense as calling a system edible because it was designed to produce edibles" (Giere 2011: 397).

[3] Douglas Adams (1979).

social organising central to organisations (Secchi and Adamsen this vol.). At a population level, spatially enabled causal loops drive internet derived technologies and ecosystems (Smart et al. this vol). In a pan-cultural domain, systems compress information (i.e., digitally formatted strings) in ways that serve collective goals. While Smart et al. (this vol.) use a functionalist model, the internet also seems to affect the lives of human subjects. In allowing such a possibility, the systemic approach invites a plurality of views. Indeed, cases like the internet and social organising matter to all who ask how people distribute control as they link routines, make instant judgements and coordinate culturally organised modes of action. Accordingly, *Beyond the Brain* unites fine and coarse grained approaches to the diversity of human thinking.

Becoming Radicalised

Our focus is empirical investigation of how body, brain and world conspire in enabling people to undertake tasks that are cognitively demanding. In retrospect, it is amazing that, even a few years ago, many doubted that this occurred: indeed, in a Psychology Department one of us felt like a heretic in arguing that problem solving drew on events from outside what philosophers called ‘the mind’. In organising the 2009 ‘distributed thinking symposium’, we invited almost everyone in the UK who took a (broadly defined) distributed view. In this section, we look at how the field has changed and, as this happened, how some of us who were there at the start became radicalised. In telling this story, we spell out a systemic view to invite *challenges* to how we conceive of human cognitive agents and, of course, their role in thinking activities.

Language-behaviour is often separated from, say, decision-making or problem solving. This echoes disciplinary divides and, indeed, the folk view that language is reducible to its salient (to us) verbal aspect. Indeed, while linguistic embodiment is widely studied (without use of the term), facts about phonetics, gesture and finely timed coordination are often used to challenge appeal to the neural representation of language types. In a well-worn slogan, language is not a code (e.g., Love 2004, Kravchenko 2007; Port 2010; Cowley 2014) but is, rather, observable, public activity. If language is seen as distributed (Cowley 2011), its neural organisation is traced to social experience with nature, artifacts and the customs of encultured people. Further, those who pursue cognition beyond the brain often rely on reports of experimental work and intuitions. For example, Clark and Chalmers (1998) make the case for ‘extended mind’ by appeal to both fictional diaries and how people consult their ‘internal’ memory. An imaginary Alzheimer’s patient, Otto, argues that, if he uses a diary to compensate for his loss by navigating the city, the artifact instantiates his beliefs. For Clark and Chalmers (1998) the diary is part of the cognitive system. While the claim has excited much debate (e.g., Rupert 2010;

Sutton 2010)[4] Clark and Chalmers (1998) have never addressed *how* Otto uses body-diary coupling. Rather, using coarse grained analysis, they stress that external resources can play a functional role akin to the body's. Today, while few dispute the fact cognition is embodied and situated, there is much debate on *how* the world beyond the brain affects thinking. Indeed, our radicalisation began in experimental work on material artifacts in problem solving. In the cheap necklace problem, for example, Fioratou and Cowley (2009) found parallels between manipulating a necklace and pico-scale linguistic embodiment (for examples, see Steffensen this vol.; Cowley 2014). We concluded that problem solving arises, not in relation to a necklace qua physical object but to the interactivity that gives it cognitive life as a thing (i.e., through its perceived affordances). Human cognitive agents experience handling that, coupled with intelligent action, brings forth emotion, observations and physical events (e.g., serendipities). Like all human behaviour, problem solving depends on action-perception or human cognitive agents who use interactivity to probe for a solution.

Using particular moments of experience human agents use language behaviour (in a pico-scale) just as they see into an X ray, for example, by saccading that follows expert gaze (Ball and Litchfield, this vol.). We thus regard human action-perception (as opposed to 'mind') as basic to extended cognitive systems: even representational 'things' are better understood through sense saturated coordination (see, Kirsh this vol.). A focus on coordination drives radicalisation. Steffensen (this vol.) posits *interactivity* as the underlying substrate of human cognition. Building on this, his team developed a qualitative method (Cognitive Event Analysis or CEA) that pursues phase shifts, changes in trajectory and lived experience. In that CEA focuses on particulars, it fits neither the axioms of enactivism (and appeal to sense-making) nor models based on experimental normalisation. Further the case is general: even in the emergency ward, people coordinate understanding by using the pico-scale of talk, action and other expression (see Pedersen 2012, 2015): they use what Malafouris calls enactive signification.

Over recent years, dynamical systems models and distributed-ecological views of language have increasingly converged with how intelligent action draws on the sensorimotor and the phenomenological. Some trace language to interactivity, others describe it as coordination, model its synergies, or emphasise that people make and track vocal tract gestures as they use facial and manual expression. Second, anti-codism extends to encultured nature (e.g., Deacon 1998, 2011) and, in Alva Noe's view of perception, make conscious experience and language part of what we do (Noe 2004, 2009). Further, Hutto and Myin (2012) challenge representationalism by denying that inner 'content' is compatible with naturalism. Even

---

[4] The hypothesis of extended mind Clark and Chalmers (1998) posits a parity principle that allows resources beyond the brain to be partly constitutive of cognition (qua function). For Sutton (2010), the question is better addressed in relation to how the external complements the psychological; for Rupert (2010), cognition is a 'phenomenon' whereby a person adverts to "a persisting, integrated set of mechanisms, capacities, abilities, etc." There are many other views: indeed, even the dual control of the PCS has a bearing on these debates.

if taking a propositional view of language (see, Harvey et al. 2016), Hutto and Myin show the viability of anti-intellectualism. Many thus adopt Chemero's methodologically motivated anti-representationalism—by default, language and cognition arise from a history of sensorimotor experience. This is consistent with the increasingly widespread view that brains co-evolved with action and perception. The resulting percaction (Berthoz 2012) can be traced to simplex principles that recur across biology and, combined with observation, grant rich culturally mediated choice-making ('vicariance'). Indeed, once language is traced to skilled action, as Bottineau (2012) shows, debate about languages and their diversity takes on new life. In a molecular scale, simplexity can help clarify the chromatin system's self-regulation (Markos et al. this vol.) and is, compatible with Maturanian views on the emergence of observers. Indeed, even neuroscience is recognising that, like culture, brains are plastic. In reading, an ontogenetically derived capacity, Dehaene (2009) uses neuronal re-cycling; building on thousands of fMRI studies and a concern with evolutionary time-scales Anderson (2014) argues that re-use is a basic principle of brain function. A long history of bodily action-perception has reshaped networks that, in many species, set off reiterated re-use of functional structures. Where, as in humans, these enable learning, this leads to increased vicariance and thus behavioural flexibility.

Interactivity

The power of interactivity was first recognised in Human-Computer-Interaction (HCI) . In this domain, people link individual histories to sense-saturated coordination as they draw on available resources. Users need little understanding. Even if 'interactivity' often designates merely an HCI focus on users (e.g., Kiousis 2002), David Kirsh (1997) traces out 'interactivity cycles'. He finds that linking action and perception with exploration enables computer users to anticipate events.[5] As people use the screen, interactivity triggers thinking. Echoing both the Soviet work inspired by Bernstein and the American empirical tradition of James and Gibson, Chemero (2013) thus shifts emphasis from input to (perceived) results. In the first edition of *Beyond the brain*, we propose, "Bidirectional dynamics connect sense-making with actions, words and feelings. As people shift between roles, human behaviour appears adaptive, dependent on learning, and, compared to other primates, flexible. Much depends on, not what a person knows, but skills in distributing control" (Cowley and Vallée-Tourangeau 2013, p. 262). Calling this *interactivity*, we stress that a cultural environment enables people/artifacts to connect "norm-based procedures with the statistical power of information compression" (Cowley and Vallee-Tourangeau 2013, p. 2). For Kirsh (1997), people can

---

[5] While this artificial kind of interactivity pertains to an individual's thinking—not processes between people—it shows how a machine can sponsor human actions.

formulate and *interpret* as, for example, putting a string like “Mongolian” into a Google search. By monitoring the screen, user actions set off output that shapes possible interpretations. Functional *results* use the PCS, as a living subject manages *projections*. For example, if an Ulan Bator circus is in town, he or she may change her plans for an evening. Unexpected results shape routines around repetition with variation. As use of interactivity becomes automatised, cultural products shape interpretative responses to things situations and events. A skilled user draws on vicariance to bring off anticipatory action (and inaction). For example, people learn how to search and which web sites to avoid! There is no doubt that the machine elicited model of interactivity is powerful. Yet, one must be wary of blind generalisation and there are arguments for developing the concept in quite different ways (see Harvey et al. 2016).

Interactivity arises in acting intelligently while monitoring results. As in solving the cheap necklace problem, solution finding uses tactility, emotion, serendipity and other means. Indeed, interactivity also raises efficiency and solution rates in, say, arithmetic problems (Vallee-Tourangeau 2013) and geometry (Kirsh this vol.). In the working world, a case of insight has been traced to how interactivity prompts ritualised ‘rejection’ of a document (Steffensen this vol.): parties to move ‘from a feeling to figuring out what to do.’ They link sense-saturated events with fears about non-payment of an invoice. The events induce functional reorganization: material objects prompt individuals to rely on invoking past acting, feeling and speaking. They recall previous occasions, use linguistic cognition, and seek out possible actions. Like slow processes of adaptation, development and learning, interactivity thus shapes skills, strategies and understanding. Indeed, this led, in 2012, to the foundation of the *International Society for the Study of Interactivity Language and Cognition* (ISSILC). In stressing the social and cultural aspect of agent-environment interaction, concern with extended systems has opened up new ways of seeking to understand cognitive human agents.

Thinking: Subjectivity in Ecosystems

Distributed cognition opens up the study of cognitive ecosystems (Hutchins 2014). Unlike extended mind or Varelian enactivism, the approach offers no central thesis or set of axioms. Rather, as an anthropologist, Hutchins chooses to trace the course of human thinking to bidirectional coupling of persons and ‘things’ in cognitive ecosystems. In cases as diverse as road construction teams (Perry this vol.), crime scene investigators (Baber this vol.) or using the internet (Smart et al. this vol), thinking draws on what Wittgenstein (1958) calls human ‘forms of life’. Values realising arises in cognitive ecosystems as living subjects undertake various projects. Aggregates of individuals and groups (both human and nonhuman) use interactivity in ways that also address subjective concerns. Timo Jarvilehto’s systemic psychology opens up a view of human agency as, at once, systemic and subjective. As in the Soviet tradition, human functionality combines the results of social and individual history: like farmers, people connect circumstances with potentials rooted in experience.

Rather than debate minds (or mind), we are concerned with ‘meaning and value’. As in Gibson’s work, we ask how human agency draws on perception-action. In the terms of Hollan et al. (2000), social processes and artifacts matter no more than do how *cultural products* affect later behaviour. However, in asking how agents deal with perturbances or, indeed, in neurocentric work, time is crucial. Thus, if human beings are partly ‘open’, historically derived or collective resources can grant a changing degree of solo control: as in material engagement theory (Malafouris 2013, 2015), the plasticity of brain and culture (or ‘metaplasticity’) allows multi-scalar cognition. Bodies individuate as living subjects as people become familiar with the cognitive ecosystems that shape skills with cultural products. Use of material symbols and smartphones is common in late modernity: they depend on literacy and, specifically, what Kravchenko (2009) aptly calls *symbolisations*. Indeed, as individuals master cultural resources, they grant human life the dual control that enables them to connect dynamics/events with objects/things in a world of situations that are inseparable from cognitive ecosystems.

Cultural ecosystems aid people in perceiving/acting. Human judgements thus depend on the *perceiving as* (Hutchins 2014) that binds interactivity to phenomenality. Given the PCS, tools and techniques can be used and/or materialised as agents mesh practices, procedures and skills. Further, where know-how matches language, it can set off in pattern seeking and even pattern making. Unknowingly, people integrate history into perception-action in when, for example, they see the constellation known as the Big Dipper or use taste to add flour making pasta. As Morris (1938) notes, such abilities extend the sensorium. The same applies in social settings where one learns to discern covariances between shifts in posture, facial expression and fluctuations in pitch, pace and timing speech pulses. Indeed, one can habituate to individual skills and capacities or, as Hodges stresses (2007), use the results to *realise values*. Similar skills appear in, say, reading X rays (Ball and Litchfield this vol.), generating narratives at a crime scene (Baber this vol.) or using technology to self-monitor (Smart et al. this vol.). Human powers are historical: in geometry, symbolisations (‘external representations’) can be used to simulate the workings of Greek logicians. The thinking that goes on can be analysed into what Kirsh calls projecting, materialising and monitoring. Interactivity links feeling, experience of geometry and cultural procedures that connect brain and world-side activity. Indeed, this is why fine grained accounts clarify how the world contributes to human cognitive agency. In extended systems, persons—living bodies with a history—adjust perception to coordinate what they and others observe. They use what Poulsgaard and Malafouris (this vol.) call the coextensive entwining of space and time: a subject’s chronotopic orientation changes thinking. To avoid such complexity, the folk reduce thinking to a process under the furrowed brow or that philosophers and psychologists often use an ‘intellectualist’ view.

Coextensive Entwining

In turning from anti-intellectualism to a positive view of cognition beyond the brain, we present the state of the art. We do so by connecting ecosystemic and person centred themes. As a result of the authors' commitments and concerns the approaches fall between weak or situated views on the internet to radical use of material engagement theory in computer aided design and rethinking the events of the Paris Commune. Next, we stress that such diversity must be described at different levels of granularity.

Turning from methodological individualism, Vallée-Tourangeau and Vallée-Tourangeau's research programme aims to compare classic cognitive findings with what the SysTM model predicts. Cautiously, they allow memory systems (i.e. semantic, episodic and procedural) to shape events in the scales of action, working memory and tracking of the world. Classic models thus supplement fine grained views of how tactility links brain-side events, observable experience and interactivity. As subjects and degree of interactivity vary, decisions are often non-optimal. On his view working memory connects 'processing' to perceived opportunities for action (SysTM ignores speech and body-based intuitions). Human agents are neither environment driven organisms nor autonomous modular systems. Rather, 'processing pathway loops' contribute to agent-environment relations. A central executive prompts attention to shift between circumstances, working memory and its long-term (brain-side) counterpart. As hands influence thinking, human flexibility ceases to be brain bound: since vocal, instrumental and full-bodied activity might all influence thinking it is possible that, later, they might be added to a SysTM model. While in early stages, the programme combines fine-grained investigation of interactivity with normalised or (coarse grained) information processing models. It thus points towards testing hypotheses about how coextensive entwining sets off results that covary (or not) with systems ascribed to the mind/brain.

In pursuing artificial networks, Smart et al. (this vol.) ask how the internet impacts on human cognition. Using tradition to take a moderate position, they treat the internet as a coding system whose users are detached from the human world. In appealing to situated cognition, they leave aside apps, modalities and media –or enactive signification. Instead, they use a metaphor of cognitive scaffolding and an organism-centred view of function. The authors show that this offers a rich description of the internet's (positive) outcomes that include, say, cityscape navigation or using 'self monitoring' to control metabolic activity. However, by reducing thinking to computation (mental states are co-constituted by extended systems), the internet becomes a shaper of habits: smartphones and computers are separated from how human subjects orient to meaning and value. The situated approach leaves no room to discuss economic, political or social consequences of an internet connected world. From this perspective, human flexibility relies on evolved responses that, broadly, link perception of available resources to how a brain has attuned to systems of electro-chemical rewards.

Madsen (this vol.) uses a systemic perspective to rethink the brain's role in human agency. He stresses that coextensive entwining depends on a temporal element that is constitutive of cognition. Acting as a self is a fluid process that uses a cultural array of techniques and technologies. Turning from neural programs or 'information processing, he asks about necessary conditions. Human cognitive agency—and questions of meaning and values—are traced to, not the brain, but co-evolution. Fluid subjectivity arises as personhood, or a fluid self, manages a delicate balancing of biological and historical concerns. Human lives thus draw heavily on co-action: on this view, rather as the internet favours global spread of information, its temporality is narrowed. People increasingly use mediated ways of perceiving, thinking and acting that, for statistical reasons, underplay history: on this view, like other cognitive practices, the web alters "beliefs, perception and memory". Aston's (this vol.) re-reading of the history of the Paris commune offers a related view. Using *Material Engagement Theory*, history is an ecosocial process where hylonoetic forces drive cognitive process. The events which Marx saw as a harbringer of proletarian revolution can—and should—be rethought as outcomes based in extended cognitive systems. On this view, and in contrast to their situated counterparts, weight shifts from information to enacted signifiers such as showing a red flag. In a socio-cognitive ecology like the city (or, by extrapolation, the internet) much depends on people who use resources made available by an energy-matter flow. Thus while Madsen emphasizes subjectivity, Aston turns to events that bring about the temporary conditions (the Commune) that prefigured new subjective ways of being.

Secchi and Adamsen (this vol.) turn from history and subjectivity to how extended systems bear on organisations and the actual world. In so doing, they find that most work which pursued cognition in organisations has taken an (input/output) view. This reduces organisations to non-cognitive legal and economic entities that depend on individual decisions, leadership styles, sense-making etc. Proposing a 'radical systemic view', they argue that, far from being boundedly rational, interactivity enables persons to constitute environments and ecosystems. They emphasise social relations, dynamics and human relationships—factors that shape organisational effectiveness. As argued by Secchi and Cowley (2016), to pursue more than efficiency, attention must fall on managing the social motor of organisational life. Most decisions result arise in social organising or doing things as people assign values by cooperating in cognitive tasks. Much depends on external resources affect ways of working—and technical supports—that sustain creative and non-creative ways of using of procedures, routines and habits. Poulsgaard and Malafouris (this vol.) offer a case study of meso-scale dynamics. Turning to computer aided design, they explore perspectives that change the world. In discussing the work of Bezier, Gaudi and Burry, they show how 'material and immaterial registers' intertwine in creating, on the one hand, the Renault and Citroen cars of the 1960s and, on the other, the glories of the Sagrada Familia (past and to come). On this view, like an organisation or a hand, a computer influences a hyloneotic field as human cognitive agents bring forth enactive signification. In material engagement theory, the process links human ingenuity to a material form of conceptual blending that drives realignment of resources in multiple time scales. By taking a distance from representationalism, human cognitive agents are able to augment what is perceived by developing a 'projective flexibility of mind'. Both conscious and unconscious intuitions use a history of experience. As mental phenomena arise with coextensive entwining, a material engagement view challenges externalism: cognition is locationally uncommitted. The radical implication is that human action-perception shapes cognitive powers—and design principles—that pertain to a domain far beyond any living body.

Changing Cognitive Ecologies

Were cognition brain-centred or based in Varelian sense-making, there would be no need for a principle of cognitive separability. However, appeal to an explanatory principle such as computation or sense making strikes us as unnecessarily narrow. Cognitive ecosystems enable human cognitive agents to use events, materiality and, above all, contingencies: they need dual control to act as social agents and living subjects. Quite simply, thinking (or thinking in action) changes with the use of resources that connect contingencies, expertise, experience with perception-action. Thoughts (qua conscious intuitions) derive from experience and, while neutrally enabled, they also rely on interactivity. Even if Kirsh's computation-based model is limited in scope, human action-perception is foundational to all kinds of *thinking*. Accordingly, cognition reaches beyond the brain and links cultural ecosystems with the subjectivity of human cognitive agents. In our view, future work will further clarify the phenomena that constitute interactivity by pursuing variability between cultures, persons, settings and moments. Much depends on, not just circumstances, but how human agents draw on expertise to attend to activity, the observed and the felt events that shape solo and cooperative action.

Cognition beyond the brain features a spirit of pluralism. While functionality is crucial to all kinds of memory, phenomena vary in granularity and, thus, demand the use of different approaches. On the one hand, some adopt Chemero's principle of tracing language, thought and action to a history of agent-environment interactions; on the other, some appeal to external representations (of content) or, indeed, inner assemblies of language-types, numbers, and rule-like operations said to be encoded by brains. Accordingly, classic computational models need to be tested alongside radical embodied models. Indeed, this is why we have stressed how the aspirations and ambitions of weak views contrast with their stronger counterparts. Thus, while functionalist tradition pursues value free models, this does not apply when values are associated with how phenomenal experience meshes with interactivity. The same contrast applies in sociopolitical domains. Does the internet always function by offering more choice and useful new habits? Or it is also having negative effects on human cognitive agency?

Human cognition is chronotopic: its spatial and temporal unity shapes how, in given circumstances, action connects with perception. Indeed, all living systems connect temporalities and, for this reason, neither brains nor metabolism need computational real time. Yet, the folk view tempts many to invoke either the experience of thoughts (conscious intuitions) and/or finished products and observable outcomes that are said to ‘reflect’ a thinking process. Appeal is made to readily operationalised temporal measures—and theories of parallel and/or linear ‘process’. All such views are simplistic: even in surfing the internet, interactivity serves to bring forth patterns on screen as computer coding alters subjective experience. There is no underlying thinking. Further, in comparison with using pen and paper to do geometry or battling to solve the invoice problem (Steffensen, this vol.), internet surfing is shallow. The user acts and perceives without recourse to the temporal complexity of subjective experience. There is nothing like simulating logicians by ‘working out’ a problem or ritualising ‘how it feels’ to receive an unpayable invoice (see, Steffensen (this vol.). Thinking activity is chronotopic and thus connects artifacts, circumstances, experience and human relations. Even if the outcomes are actual, the resources of reason place thinking neither ‘in’ the head nor ‘in’ the world. Human cognition depends on using routines, rules, words, numbers, bodies of knowledge and is, for this reason, non-localisable. Given the importance of cultural products, people-in-cognitive-ecosystems can use circumstances to make subjective judgements—even in writing/reading a paper. Although usually performed on one’s own, the *thinking that goes on* is social—and, in spite of this, shapes a flow of self-generated subjective experience.

References

Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain*. Cambridge, MA: MIT Press.

Adams, D. (1979). *The hitchhiker’s guide to the galaxy*. London: Pan Books.

Aston, A. (this vol.). Unruly minds in the city of light: Cognitive ecology and the Paris Commune of 1871. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Baber, C. (this vol.). Distributed cognition at the crime scene. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Ball, L. J., & Lichfield, D. (this vol.). Interactivity and embodied cues in problem solving and insight: Further contributions to a “theory of hints”. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Ben Naim, J., Bonnefon, J. F., Herzig, A., Leblois, S., & Lorini, E. (this vol.). Computer-mediated trust in self-interested expert recommendation. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), Cognition beyond the brain: Interactivity, computation and human artifice (2nd ed.). Dordrecht: Springer.

Berthoz, A. (2012). *On simplexity: Simplifying principles for a complex world*. New Haven: Yale University Press.

Bottineau, D. (2012). Remembering voice past: Languaging as an embodied interactive cognitive technique. In: Proceedings of the Humanities International Forum, Russian St*ate University for the Humanities*, Moscow.

Chemero, A. (2013). Radical embodied cognitive science. *Review of General Psychology*, *17*, 145–150.

Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis, 58,* 7–19.

Cowley, S. J. (2011). *Distributed language*. Amsterdam: Benjamins.

Cowley, S. J. (2014). Linguistic embodiment and verbal constraints: Human cognition and the scales of time. *Frontiers in Psychology, 5,* 1085. doi:10.3389/fpsyg.2014.01085.

Cowley, S. J., & Vallée-Tourangeau, F. (2010). Thinking in action. *AI & SOCIETY, 25,* 469–475.

Cowley, S. J. & Vallée-Tourangeau, F. (2013). *Systemic cognition: Human artifice in life and language.* In S. J. Cowley, and F. Vallée-Tourangeau, F. (Eds.), *cognition beyond the brain: Computation, interactivity and human artifice*, pp. 255–273. Dordrecht: Springer.

Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. WW Norton & Company.

Deacon, T. W. (2011). *Incomplete nature: How mind emerged from matter*. WW Norton & Company.

Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. Harmondsworth: Penguin.

Fioratou, E., & Cowley, S. J. (2009). Insightful thinking: Cognitive dynamics and material artifacts. *Pragmatics & Cognition, 17,* 549–572.

Giere, R. N. (2004). The problem of agency in scientific distributed cognitive systems. *Journal of Cognition and Culture, 4,* 759–774.

Giere, R. N. (2011). Distributed cognition as human centered although not human bound: Reply to vaesen. *Social Epistemology: A Journal of Knowledge, Culture and Policy, 25,* 393–399.

Harvey, M. I., Gahrn-Andersen, R., & Steffensen, S. V. (2016). Interactivity and enaction in human cognition. *Constructivist Foundations*, *11*(2).

Hodges, B. H. (2007). Good prospects: Ecological and social perspectives on conforming, creating, and caring in conversation. *Language Sciences, 29*(5), 584–604.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction, 7,* 174–196.

Hutchins, E. (2014). The cultural ecosystem of human cognition. *Philosophical Psychology, 27*(1), 34–49.

Hutto, D. D., & Myin, E. (2012). *Radicalizing enactivism: Basic minds without content*. MIT Press.

Jaynes, J. (2000). *The origin of consciousness in the breakdown of the bicameral mind*. Houghton Mifflin Harcourt.

Järvilehto, T. (1998). The theory of the organism-environment system: I. Description of the theory. *Integrative Physiological and Behavioral Science, 33,* 321–334.

Järvilehto, T. (2009). The theory of the organism-environment system as a basis of experimental work in psychology. *Ecological Psychology*, *21*, 112–120.

Jones, P. E. (2013). You want a piece of me? Paying your dues and getting your due in a distributed world. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (pp. 113–129). Dordrecht: Springer.

Kiousis, S. (2002). Interactivity: A concept explication. *New Media and Society, 4,* 355–383.

Kirsh, D. (1997). Interactivity and multimedia interfaces. *Instructional Science, 25*(2), 79–96.

Kirsh, D. (this vol.). Thinking with external representations. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Kravchenko, A. (2007). Essential properties of language: Why language is not a digital code. *Language Sciences, 29*, 650–621.

Kravchenko, A. V. (2009). The experiential basis of speech and writing as different cognitive domains. *Pragmatics & Cognition, 17*(3), 527–548.

Love, N. (2004). Cognition and the language myth. *Language Sciences, 26,* 525–544.

Madsen, J. K. (this vol.). Time during time: Multi-scalar temporal cognition. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Malafouris, L. (2013). *How things shape the mind: A theory of material engagement*. London: MIT Press.

Malafouris, L. (2015). Metaplasticity and the primacy of material engagement. *Time and Mind: The Journal of Archaeology, Consciousness and Culture, 8,* 351–371.

Markoš, A. Švorcová, J. & Lhotský, (this vol.). Within the skin—and beyond. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Morris, C. W. (1938). *Foundations of the theory of signs*. Chicago: Chicago University Press.

Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. Macmillan.

Noë, A. (2004). *Action in perception*. Cambridge MA: MIT press.

Neumann, M. and Cowley, S. J. (this vol.). Agency, languaging and the resources of reason. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd Edition). Dordrecht: Springer.

Pedersen, S. B. (2012). Interactivity in health care: Bodies, values and dynamics. *Language Sciences, 34*(5), 532–542.

Pedersen, S. B. (2015). *The cognitive ecology of human errors in emergency medicine: An interactivity-based approach*. Unpublished PhD dissertation, University of Southern Denmark.

Perry, M. (this vol). Socially distributed cognition in loosely coupled systems. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd Edition). Dordrecht: Springer.

Poulsgaard, K. S., & Malafouris, L. (this vol.). Models, mathematics and materials in digital architecture: Creative imagination at the human-computer interface. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Port, R. F. (2010). Rich memory and distributed phonology. *Language Sciences, 32*(1), 43–55.

Ross, D. (2007). H. sapiens as ecologically special: What does language contribute? *Language Sciences*, 29/5: 710–731. R. D. Rupert. (2009). *Cognitive Systems and the Extended Mind*, Oxford: Oxford University Press.

Rupert, R. D. (2010). Extended cognition and the priority of cognitive systems. *Cognitive Systems Research, 11*(4), 343–356.

Secchi, D., & Cowley, S. (2016). Organisational cognition. In *Euram Conference Proceedings*.

Secchi, D., & Adamsen, B. (this vol.). Organisational cognition: A step too close. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Smart, P., Heersmink, R., & Clowes, R. (this vol.). The cognitive ecology of the web. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Steffensen, S. V. (this vol.). Human interactivity: Problem-finding, problem-solving, and verbal patterns in the wild. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Sutton, J. (2010). Exograms and interdisciplinarity: History, the extended mind, and the civilizing process. In R. Menary (Ed.), *The extended mind* (pp. 189–225). Cambridge MA: MIT Press.

Vallée-Tourangeau, F., & Villejoubert, G. (2013). Naturalising problem solving. In *cognition beyond the brain* (pp. 241–253). London: Springer.

Vallée-Tourangeau, G. & Vallée-Tourangeau, F. (this vol.). Cognition beyond the classical information processing model: Cognitive interactivity and the systemic thinking model (SysTM). In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice* (2nd ed.). Dordrecht: Springer.

Winnicott, D. W. (1969). The use of an object. *The International Journal of Psycho-Analysis, 50,* 711–716.

Wittgenstein, L. W. (1958). *Philosophical investigations* (2nd ed.). Oxford: Blackwell.

Chapter 2
Socially Distributed Cognition in Loosely Coupled Systems

Mark Perry

Abstract Distributed cognition provides a theoretical framework for the analysis of data from socio-technical systems within a problem-solving framework. While the approach has been applied in tightly-constrained activity domains, composed of well-structured problems and highly organised infrastructures, little is known about its use in other forms of activity systems. In this paper, we explore how distributed cognition could be applied in less well constrained settings, with ill-structured problems and loosely organised resources sets, critically reflecting on this using data from a field study. The findings suggest that the use of distributed cognition in an augmented form can be useful in the analysis of a wide range of activity systems in loosely coupled settings.

The question of how thought is distributed over a variety of non-neurological resources has received a considerable amount of interest, both within the communities of cognitive science and beyond to anthropology, sociology and even computer science. Yet while theoretical frameworks are being developed and tested to identify the socio-cognitive mechanisms through which this externalised form of cognition is organised and co-ordinated, this body of empirical research is typically limited to relatively narrowly bounded systems of information representations and agents. At the same time, the sorts of problems that it has examined are tightly defined (i.e., there are clearly articulated and specific goals), so that the problem solving agents are aware of what the state of the final problem resolution will be. This is important and groundbreaking work for understanding human behaviour and thinking ‘in action’. However, many problem-solving situations are not so well structured and resource-delimited. This includes settings in which problem solving is distributed over technical, social and organisational systems that do not have well defined boundaries, so that multiple computational structures may be employed, or where the problem solving resources available to these systems may dynamically

---

M. Perry (✉)
Department of Computer Science, Brunel University,
Middlesex, Uxbridge UB8 3PH, UK
e-mail: Mark.Perry@brunel.ac.uk

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_2

change: agents can co-opt a range of tools into their work and new agents may be introduced into the system. Similarly, there are problem-solving situations in which the problem is itself not clearly defined or underspecified (these are known as ill defined or ill structured problems (Simon 1973)), and part of the problem solving activity is to determine the form of the solution. We call these 'loosely coupled' systems.

The importance of loosely coupled systems is that are they are typical of many kinds of problem solving behaviour seen in socially organised collective activity and organisational life. Understanding problem solving within these system is therefore a matter of great practical concern for their designers and managers. However, and as we will demonstrate in the paper, the analytic techniques that have been developed to understand problem solving in tightly delineated systems undertaking well specified activities are not easily tuned to these loosely coupled systems. Deriving an analysis of such activity directly from the methods used in existing studies may therefore prove to be problematic.

One of the most clearly articulated approaches used to examine distributed problem solving that extends cognition into artefacts and social relationships is known as distributed cognition (DCog; Hutchins 1995a, b). With some variations in its application (e.g., Wright et al. 2000; Blandford and Furniss 2006), it is perhaps the most clearly articulated, critiqued, commonly used and well known form of exploring how distributed action can be examined as a cognitive process (see, for example, Giere 2004), and it is for this reason that we have chosen to re-examine its use here. Distributed cognition draws its theoretical and analytical foundations from cognitive science, augmenting these with techniques for primary data collection and ethnographic representation drawn from sociology and anthropology. It is unusual in that it seeks to study people, social interaction and artefact use within a common framework. The focus of DCog is on the representational transformations that occur in external media. In practice, information about the settings is usually captured using ethnographically informed observations and interview-based methods. This has been described as 'cognitive ethnography' (Hutchins 1995a). Here, attention is concentrated on the agents (who co-ordinate the representational transformations) and the artefacts in the environment (media that encode the physical representations of the system's internal state). The interaction between these elements is crucial in understanding how representational transformations are co-ordinated. We distinguish this 'flavour' of distributed cognition developed by Hutchins and drawing from cognitive science as DCog, and differentiate it from other uses of the term. As noted by Halverson (2002), there are problematic issues in the various interpretations of the term 'distributed cognition', which is used in an analytically different way by writers such as Heath and Luff (1991), and in the book 'Distributed cognitions' (Salomon 1993), or simply as a synonym for socially shared cognition (see Engestrom and Middleton 1996; Resnick et al. 1991).

While DCog has been successfully applied in understanding some working environments, its application has been limited to a tightly constrained set of problems, with a pre-specified set of active agents and a limited set of representation bearing artefacts. As McGarry (2005, p. 187) notes, these analyses have taken place in 'highly rationalised environments' precisely because the idea of 'social systems as computation' maps most closely onto these types of setting as a solidified cognitive architecture with an easily identifiable end-point. As it stands, DCog has proved a useful and insightful approach to examining these settings, and the focus of the approach (as applied in these studies) matches well with the structured settings of this work. Yet, as we have noted, characteristics such as these are not shared in a vast range of work settings, such as team-based game playing, design, financial services, or medical work. It is interesting to note here that the existing DCog analyses have focused on partial elements of wider work settings (e.g., in navigation, Hutchins 1995a; air-traffic control, Hutchins 1995b; Halverson 1995; fMRI brain mapping, Alač and Hutchins 2004), emergency control centres (Artman and Waern 1999) and telephone call centres (Halverson 2002). The nearest that DCog analyses have come to this in their application to highly dynamic and loosely structured settings are Hazelhurst's (1994) and Holder's (1999) Ph.D. theses, although these do not critically examine the methodological extensions that have been made to extend the approach. Moreover, the problems described within even these settings have been carefully 'bounded' to create a simplified problem space that does not account for the larger setting.

As an example of this, we put forward ship navigation as the archetypal complex and socially supported problem that DCog analyses have sought to examine. Drawing from Hutchins's (1995a) study of navigation on a US naval vessel this is, in his own words, 'a partially closed information loop' (p. 39):

> The activities of the Navigation Department revolve around a computational ritual known as the *fix cycle*. The fix cycle has two major parts: determining the present position of the ship and projecting its future position. The fix cycle gathers various bits of information about the ship's location in the world and brings them together in a representation of the ship's position. (ibid., p. 26)

This process involves taking visual bearings that can be translated into lines that should intersect on a navigational chart. At the risk of simplifying such a complex operation, this begins when two 'pelorus operators' use a telescope-like device (alidades) to take visual bearings of landmarks, and assign a precise directionality to these bearings ('lines of position'). This information is communicated to the bearing recorder and plotter in the chartroom who mark this on the chart as a pencil line using a 'hoey' (a one-armed protractor) to map the lines of position accurately onto the chart. All of these tasks need to be done on a highly fixed timeline, in the correct order, and parts of this can only be undertaken by particular people. Various technologies are employed in the fix cycle (e.g., charts, pelorus, hoey) and it involves six team members performing various specialised tasks. Where breakdowns occur in the standardized navigational process, 'the team are able to compensate for local breakdowns by going beyond the normative procedures to make sure that representational states propagate when and where they should' (Hutchins 1995a, p. 228). Nevertheless, despite micro-scale adjustments in the tools used and organizational arrangement between the system's components, the end result of the fix cycle is the same (a set of lines of position on a chart), and the group is unable (or at least in the analyses presented) to enrol additional resources outside those that already form the existing set.

Limiting this scope, of course, has benefits in providing a focus to the analysis. By artificially setting the problem situation up as a ‘closed’ system, the DCog analyst does not need to face the problems of extending his or her analysis outside of these limited settings. While the analyses that existing studies have developed are extremely valuable, extending the domain of interest beyond the limited systems explored would provide an interesting and valuable insight into the ways that these sub-systems were embedded into the larger work systems of the workplace. So, developing Hutchins’ (1995a) work on navigation for example, we might learn how navigational information is interpreted in the other functions of the ship, and how other aspects of activity on the ship might impact back into the navigational sub-system. This type of analysis poses a different set of problems to those examined in these previous studies, but such analyses would nevertheless be interesting and useful to understand wider systems of context activity than the simple(r) activity in focus. Indeed, Hutchins does touch on this as a practical concern, stating that ‘low-level functional systems may be embedded in larger functional systems’ (p. 153), although in the cases described, even these larger functional systems as described are fairly restricted in their complexity and reach.

In the following sections, the foundations and existing use of DCog will be discussed and examined to highlight the major concerns over its use in large and complex organisational settings. Briefly reviewing the theoretical and methodological underpinnings of the DCog framework will allow us to identify and discuss the areas of difference between tightly and loosely coupled systems. To demonstrate this, we then draw on a field study, pulling out findings to highlight the application of DCog and to motivate suggestions for the development of DCog and its evolution as an analytical method.

Distributed and Socially Distributed Cognition

DCog in Cognitive Science. The theoretical framing and terminology used in DCog reflect its roots in cognitive science. Classical cognitive science provides a conceptual framework with which to examine intelligence and problem solving, exploring how information is represented in the cognitive system, and how these representations are transformed, combined and propagated through the system in goal-oriented behaviour (Simon 1981). These cognitive processes and representations are normally considered to be internal to the individual and held mentally. In DCog, a larger granularity of analysis is selected than that of the individual, to include the use of information artefacts in the cognitive process (see, inter alia, Hutchins and Klausen 1991; Norman 1993; Hutchins 1995b). These systems include artefacts such as a pen and paper, which can be used as memory aids, as well as more complex artefacts that are integral to the computation, such as slide rules. As such, DCog provides a coherent framework with which to structure the analysis of activity systems in terms of their informational content and problem solving activities. As well as incorporating artefacts in the cognitive analysis, DCog also differs from classical cognitive science in its views of the cognitive process. DCog suggests that the cognitive process can be described as being distributed over a number of people co-operating through social mechanisms, often referred to as 'socially distributed cognition' (Hutchins 1995a; Hollan et al. 2001).

The unit of analysis in DCog may consist of any number of representations, embodied in people, computerised artefacts and non-technological artefacts (Rogers and Ellis 1994). As a framework, DCog provides a unique insight into how technology and the socially generated media of communication act upon and transform representations, and in doing so, perform computations, or information processing activity. The aim of DCog is therefore to understand how intelligence is manifested *at the systems level* and not the individual cognitive level (Hutchins 1995a). The mix of people and artefacts in a given situation contribute to the *functional system* of activity, which includes all of the representation-carrying and representation-transforming entities involved in the problem solving activity. A distributed cognitive analysis examines the means by which the functional system is organised to perform problem solving. From a computational perspective, the functional system takes in inputs (representations), and transforms these representations by propagating them around the units of the system. A distributed cognitive analysis involves deriving the *external* symbol system (cf. Newell and Simon 1972) by capturing the elements of processing (representations and processes) that transform system inputs into outputs for particular tasks. In many cases, these distributed representations and processes are brought together by agents—people—and co-ordinated through social mechanisms.

The Division of Labour. Socially distributed cognition allows the analyst to describe group activity as a computation realised through the creation, transformation and propagation of representational states (Hutchins 1995a; cf. Simon 1981). By bringing representations in the system into co-ordination with each other, information can be propagated through the larger cognitive system, being continually modified and processed by a number of individuals and artefacts, until a desired result is reached. However, while processing of the information available to the group is analogous to an individual's cognitive capabilities, the architecture of this activity differs significantly. How these socially embodied activities are organised with respect to one another is known as the 'division of labour'. Understanding how this division of labour operates within the functional system provides a means of understanding its internal organisation and working practices—in effect it determines a large component of the *computational architecture* of the distributed cognitive system. Of course, there are many ways that work *can* be done, and this division of labour is malleable and subject to change as the agents within the functional system learn, change roles or introduce new technologies into their work.

Distributing work across a group of agents must involve some form of organisation to co-ordinate activity to develop a working division of labour, or an architecture for computationally acting on the problem representation(s) to affect a solution. To solve a problem collaboratively, this division of labour must operate so that work is broken into parts but managed so that these (partially resolved) components can be re-incorporated. Internal cognitive (i.e., mental) factors that are effectively invisible to the group (and to the analyst) need to be trained on the problem by individuals to bring their expertise to bear on these sub-tasks before they are able to re-incorporate their sub-task with the global task. This means that within the distributed cognitive system, problem-solving expertise lies not only in the knowledge and skills within individuals, but in the organisation of those individuals, and on their ability to co-ordinate their work with one another. This larger unit of knowledge organisation may be determined explicitly in predetermined protocols and procedures, or more informally through the context of their work environment being made visible or through the configuration of the artefacts used in the environment. DCog thus presents a method of describing and analysing how this overall system operates, using a computational metaphor applied to the functional system as a whole.

The extent to which the division of labour is clear-cut and explicit (i.e., visible to the analyst) in the work activity will determine the extent to which analysts can develop a description of the distributed cognitive system. Where the activities, agents and information artefacts in a workplace system are constrained and work roles are clearly specified in known protocols, the application of DCog to analyse the functional system should be relatively straightforward. This is likely to be particularly the case if the problem representation is explicit and can be tracked through its transformations in the system. Unfortunately, as any experienced ethnographer will recount, this is rarely the case, even if work with systems appears to be well defined at first impression. In many cases, the only way to find systems that are even remotely well defined is to pick out (relatively) simple sub-activities from workplaces and to examine these—as has been the case for the majority of DCog studies to date.

DCog: Application in Data Collection and Analysis. As an analytic framework, DCog focuses the analysis on the salient points relating to the cognitive (i.e., information processing) characteristics of the functional system to structure the data available. Thus, data collection within the analytic framework of DCog must allow the identification of the representations and processes in the goal-directed behaviour. The analyst needs to show how these elements are used as resources in information processing by demonstrating how this is mediated through *action*. Emphasis is placed on the role of the artefacts carrying information representations and on collaboration around these artefacts. Analysis therefore centres on the artefacts that are used or created, how they are used, who they are used by, how changes are made to them, and how the organisation structures access to these artefacts. Amongst other tasks, this may involve mapping out the information flows through the organisational structure, identifying the sources and sinks of this information, the tools used to manipulate and transmit it, and the 'chains of command' initiating activities.

While the framework of DCog is non-prescriptive in its application to data collection, one method—ethnography—has come to be pre-eminent. Ethnography has its own traditions within sociology and anthropology and has long been used as a technique for gathering naturalistic data about activity within workplace and other settings. However, at its core, ethnographic analysis provides a means of exploring *how* human activity is organised (Hughes et al. 1992; see also Hammersley and Atkinson 1995 for an overview). It allows the analyst to physically 'enter' the functional system and to build up a picture of activity from the member perspectives, but which encompasses a higher-level perspective than the individual, showing interdependencies within the division of labour. These may be invisible to the participants themselves.

In principle, the analytical framework offered by DCog is clearly of value in its analysis of loosely coupled systems, through demonstrating the mechanisms of co-ordination and collaboration in problem solving. Its focus on the 'social' aspects of information processing identifies it clearly as different from almost all other forms of analysis as it uses the same 'language' to describe all of the system components and their operation.

Comparative Analysis: Examples from a Loosely Coupled System

This section introduces a field study from which several examples are drawn, and provides a basis from which we argue for the extension of DCog as an analytic approach to the study of loosely coupled systems. Its intention is to demonstrate (i) the similarities and differences between the traditional domain of study for DCog and less constrained, loosely coupled organisational settings; and (ii) an approach to gathering data that allows us to undertake an analysis of aspects of activity within these settings.

Background to the Study Setting. The fieldwork was conducted as part of a larger project examining the use of information technology in construction, and was conducted over the course of around a year by the author who had no previous knowledge of construction. A number of field visits of two-week duration were taken over this time, as the author was embedded into the teams. This was largely non intrusive, as team members were used to apprentices and other staff shadowing them. The field study involved an examination of design and construction on a large road-building project (see also Perry 2003) that took place over a three-year period. Fieldwork covered the participation of five (fluctuating in number) spatially distributed teams employed by a construction company. To build the permanent road structures, the construction company had to design non-permanent, supporting structures known as 'temporary works' (concrete moulds, supply roads, and so on). Temporary works are derived from the designs of the permanent works, but their designers have to take into account the site conditions (slope, weather, soil condition, and existing structures) and the available resources (money, time, materials, equipment and labour) that are not documented in the permanent works designs.

Several people were involved in the functional system developing the temporary works design, including a construction team, made up of a team leader, various engineers, foremen and gangers (work supervisors), and general labour. The team operated on the site, but were supported off-site in specifying the design of the temporary works by a design co-ordinator who worked at another location on the building site. The design co-ordinator worked closely with a design engineer (located several miles away from the site) to draw up a temporary works design. A number of other individuals and stakeholder groups external to the organisation were also closely involved in the process, including a 'resident engineer' (RE), who checked that the road was designed and built to contractual standards.

The temporary works design processes were partially prescribed in the handbooks and manuals that determined relationships between people, and in the organisational structures they inhabited. These specified where responsibilities for tasks lay and determined the roles that the individuals had to fulfil. While these were generally followed, particularly the safety-related rules, they were used more as organising principles, followed when appropriate, but worked around when other methods might be more appropriate in the circumstances. Structures in the environment at the site and office also played a role in determining the processes that would be applied to particular problems. These determined the configurations of representations and processes that *could* be applied to the design problem. For example, the physical size of the site made locating people difficult, and communications were complicated accordingly. However, when in the construction team's on-site office, the engineers and foremen shared a physical space and had a wide range of media with which to communicate.

In the terminology of DCog, the transformational activity in temporary works design involved taking inputs into the functional system, transforming them into representations and re-representing them until they match the desired goals. However, the different resources available in particular circumstances, and the different ways in which individuals could resolve the problem-solving situation meant that problem resolution could be achieved in a variety of ways. Hutchins (1995a) also describes this, showing how 'local functional systems' are built up around—and exploit—the particularities of individual problem situations.

Illustrative Examples from Fieldwork

Having established the broad stetting of the field study, we can now look at it in relation to parts i. and ii. above, exploring the differences between traditional studies of DCog and the construction setting, as well as the approaches used to gather data in undertaking an analysis of aspects of construction work within the DCog framework.

Access to Resources. The key resources in design and construction are rarely pre-determined before the problem solving activity is initiated. The fieldwork illustrates how the problem solving activity changed as additional agents and artefacts become available. When new designs or requests for information arrived at the site, there were a variety of different follow-on activities that might result from this, although it was by no means ‘programmatic’ what these might be. Hutchins (1995a) recognises this ‘availability of data’ as a key controlling factor in how a functional system can organise its computational and social structure: additional information can lead to new social arrangements, which in turn lead to new computational structures (Hutchins 1995a, p. 342). Where Hutchins’s analysis diverges from the concerns relevant to our fieldwork is that these new computational structures arise from the time and frequency of fix cycle data that is available and the agents available for acting on it, while in construction, new computational structures are far more substantive and may include the arrival of a different form of (relevant) information and/or new agents with entirely new roles and responsibilities into the system. This is a fundamental difference with wide reaching consequences: it is not simply the local configuration of the functional system that can change over time, but its very constitution.

This can be illustrated with our fieldwork. The construction team’s office was an important resource in their work. It had an ‘open plan’ layout, and the engineers and quantity surveyors were able to see who else was present. It allowed them to speak to each other without moving from their desks, to overhear other people on the telephone or when speaking to each other, and to see the information laid out on each other’s desks. While the construction team were centred in this office, individuals spent a large amount of their time on-site, and the distributed nature of the construction site made contacting these individuals difficult. When people were not present to talk to directly, other media were used to communicate, either synchronously through the use of the radio link, or asynchronously, through placing written notes, sketches, method statements or risk assessments on people’s desks, or jotting notes onto a whiteboard. Messages were also left with people who were in the office for when that person came back. A photograph of the office space, its contents and layout is shown in Fig. 2.1.

As can be observed, the workplace was informationally very rich. Paper covered almost every surface, often several layers deep, even pinned onto the walls. When information was required from a person who was not physically present, this ‘desk litter’ could provide clues to their location, in the forms of the drawings and other representations on the desk, indicating the task that they were currently engaged in, and providing a guide to their current location. Other artefacts also provided information about the whereabouts of people: if a person’s Wellington boots and hard hat were missing, they were probably out on site; if someone had a pair of muddy boots under their desk, it meant that they had been on the site and could be asked about the current work situation. Depending on the weather, it was even possible to see how long ago a person had been out on the site, for example from the wet or dried mud on boots, which could be useful if one of the team was trying to locate another individual out on the site. Equipment such as the geodometer was also useful in this way—if it was missing from the office, then a graduate engineer would probably be out on site with it and could be asked to run a favour over the radio. Even the window was used to see whether a person’s vehicle was in the car

Fig. 2.1 The construction team office

park outside the office: if this was the case, then that person was highly likely to be somewhere on the site.

Spoken communication was conducted from the desks, allowing all of the participants in the room to be aware of developments, or allowing them to contribute to the discussion. When the senior or site engineers wanted to speak to the graduate engineers, they would stand up and chat over the tops of the partitions, providing a visual and auditory focus of attention in the room. This allowed people to work while keeping an ear to the conversation and keeping abreast of developments, to ask questions, but also allowing them to enter the conversation and add to the discussion. Within the literature of workplace studies, such observations are relatively commonplace (e.g., Heath and Luff 1991; Murray 1993). However, within the DCog framework, the analyst must draw from his or her observations how particular representations are propagated and transformed. Here, in the example above, the potential range of representations that participants could select from was vast, and they could be used in a variety of ways to achieve the same goal. In the event of a similar situation arising, it was unlikely that the same resources would be available in the same configurations as before, and used in the same way. This differs significantly from the examples of DCog documented in aircraft and navigational settings, where stable behavioural patterns (Hutchins calls these 'computational sequences') tend to be recurrent.

A consequence of this resource flexibility is that it is harder to build generalisations or models about problem-solving activity. In these situations, ethnographic descriptions of activity are likely to represent 'one-off' solutions, generated by members who generate and maintain a computational structure that utilises only a subset of the potential resources available. Such descriptions will therefore be useful to illustrate how and which resources are used, but they cannot be seen as anything more than exemplars, or instances of problem solving. This is informative in understanding the activity *in general*, but not so much the predictive performance of *an* activity. So, for example, it may be useful in understanding the formalisms within engineering drawings that allow people to communicate using them, but not in how a particular problem might be resolved in which engineering drawings were a component part.

Problem structure. The problems faced in construction were often poorly understood prior to the initiation of the problem solving activity (i.e., they were ill-structured), and a component of this activity was to understand exactly what the problem was as well as how to resolve it. For example, much of the physical component of construction work was demand-led, and work could only occur when the site had been prepared: materials or other resources might have to be ordered or cancelled at the last minute because the site was prepared for the work earlier or later than expected. The use of different construction methods or materials (arising from product non-availability or particularities of the situation) in the permanent structures could thus change the project's specifications. This differs from previous studies of DCog in navigation and cockpit activity, in which relatively regularised and well-structured problems are encountered: thus there is the fix cycle in navigation and the landing sequence in an aircraft cockpit. Deviations can and do occur, but these are relatively highly constrained and delineated within a well understood set of criteria—the goal is clear, and the set of operations that the participants have to operate on is limited and practiced (although see Hutchins 1991). The examples given below demonstrate the structure of the problem faced by the designers at the construction site. The section begins with a description of the global design situation and shows how that within this, ongoing problems were identified and transformed from ill-structured problems to well-structured problems.

In official project procedures, the team's senior engineer should formally present the initial 'design brief' (a temporary works design specification) to the design co-ordinator before a temporary works design could be generated. In practice, this design brief was often little more than a few ideas sketched or jotted onto a scrap sheet of paper. This occurred because the senior engineer had too little time to perform the task, and often had very little understanding of what information the design engineer might require in the problem specification. Through discussions with the design co-ordinator, a detailed specification would be generated, containing information about the site conditions, the materials, labour and other resources available to construct the temporary works structure. The construction team's senior engineer and the design co-ordinator would then pore over the permanent works drawings, the initial temporary works design brief and several sheets of blank paper. Here's an example:

> Senior engineer (SE): 'If you look here, there's a barrel run there' (points at sketch generated in the meeting of a section view through a design structure)
>
> Design co-ordinator: 'Yes I see'.
>
> SE: 'So if we dig here…' (he holds one hand to the sketch and runs a finger on the other hand along a permanent works drawing (plan view) beside the sketch, indicating a line of reference)

> Design co-ordinator: ‘No you can’t do that because of drainage problems…’ (pauses) ‘… No, no, I see now’.
>
> SE: ‘So if we cap these piles here…’ (indicates several points on the sketch)
>
> Design co-ordinator: ‘Yeah. OK. Let’s do that’.

The example shows how little was understood about the problem before commencing the design process. The problem itself (generating an effective way to provide structural support) arises out of the comparison of artefacts (sketch and drawing). Having recognised the problem, the two later went on to generate a solution. This solution is very different to the structured ways that the chart was used in Hutchins’ description of navigation. Here, the structure of the problem in the work of construction is not fully specified, and the problem solvers must endeavour to clarify what they need to do to achieve their goal (in the case of the last example, to provide an appropriate form of structural support). Typical cockpit and navigational examples in the DCog literature do not involve this form of behaviour: the problem solvers already have a known problem and a well-practised set of procedures with which to generate a solution. The importance of this difference in problem solving for DCog in construction is that the problem space constantly changes over the time period, and repetition of activities is consequently infrequent.

The fieldwork also shows a substantial part of collaborative problem solving involves the system self-organising itself. While self-organisation is recognised and discussed by Hutchins as a feature of activity within functional systems in navigation (e.g., Hutchins 1991), we have observed extensive ongoing self-organisation in our research into loosely coupled systems. A crucial component of self-organisation is the agents’ awareness about the state of the functional system, so they are able act in an appropriate and timely way to the ongoing situation. An analysis therefore needs to clearly show how agents manage to achieve situation awareness. However, situation awareness is not normally assumed to be a part of problem solving in DCog (because it is not always directly associated with a representation transforming activity) and as such has been largely ignored. We argue here, that in DCog studies of loosely coupled systems, situation awareness *must* be considered as a core feature of computational activity, even when it is not directly associated with *a particular* problem-solving event.

Organisational Structure and Problem Solving Dynamics. In this section we show examples of how the components of the functional system were organised in the construction project. In the terms used by Hutchins (1995a), there were a number of documents that determined a ‘standard operating procedure’ (or SOP) including manuals and handbooks. However, their application was not as heavily enforced as navigational or aircraft cockpit systems, for a range of reasons (in navigation, SOPs control time criticality, safety, number of participants, and personal accountability). In construction, the organisational structure inherent in the SOP provided a basic structure, but allowed flexible interpretation by the actors involved. The example below shows this in an instance of how materials were ordered by the construction team and the resourceful way in which they managed to accomplish this (from fieldwork). To set the scene, a site engineer was discussing a concrete pour with a remote person over the telephone (note: only one side of the telephone conversation could be monitored by the fieldworker):

> Site engineer: (stands up and speaks loudly into telephone) 'So, what I'm asking is: should we put concrete into the tower?' (raises his head and looks at the senior engineer with raised eyebrows)
>
> Senior engineer: 'Yes'.
>
> (Site engineer, completes the telephone call, then lifts a radio to speak to a foreman to give the go ahead. A graduate engineer overhears this:)
>
> Graduate engineer: (orients towards senior engineer) 'Do you have any spare…(pause)… can I have three cubic metres?'
>
> Senior engineer: 'OK. Yeah.'
>
> (Site engineer overhears this and radios through to the foreman to arrange it).

In this observation, the open-plan office space allowed the overhearing of telephone conversations, and was used by the site engineer as a means of asking the senior engineer if he could go ahead with construction. This was not pre-planned, but arose from a request for information arising from a distant third party. A graduate engineer, in turn, overheard this, and made a request for materials, which was arranged by the site engineer. This saved money (and effort) by ordering one and not two separate concrete deliveries, yet none of this was planned in advance. This situation was able to take place because of the open-plan structure of the office space, but also because the participants knew that they did not have to order material through the specified SOP by ordering each load of concrete through a formal materials order. In this case, a separate order of concrete would have had to be made, tying up resources and losing the economies of scale that come with a large order.

A second example shows how the organisational structure acted as a *resource* for problem solving, while the mechanisms used in resolving the problem were socially mediated and negotiated between people. In the construction project, tasks were allocated to people through a number of mechanisms, dependant on hierarchical structures of seniority and the contextually dependent features of the setting. While allowing a degree of autonomy, the participants understood their responsibilities and the roles that they were expected to perform. The example below illustrates how knowledge distribution occurred through a variety of agents and media.

A graduate engineer was asked to check on the particular characteristics of a concrete mould (known as 'shuttering') by the clerk of works (who was employed by the RE). According to company regulations, queries raised by the RE or their staff should involve recording the problem, finding the answer, and filling out a 'works record', which would be sent to the site office, placed on file, and a copy sent on to the RE. Accordingly, the graduate engineer filled out a works record form with the problem request and sketched a diagram of the concrete shuttering and the setting it was placed in. He telephoned (someone) off-site, and discovered that the information he needed about using the shuttering was in an advertising/promotional leaflet sent out by the shuttering company, and which had just been sent on to be held on file in the team's office. Almost immediately, he noticed that this leaflet was lying on one of the foremen's desks, as he had been looking through it with an eye to ordering more materials. The engineer read off the technical details from a table on the leaflet and added this information to the form. He then posted the works record to the site office for circulation. As the works record was an official form, no accompanying contextual information was required because the nature of this structured document meant that it would always be processed in the same way. Due to the slow speed of the internal postal service, the engineer later went back on site, located the clerk of works and reported his findings personally.

The example demonstrates how the members of the local functional system (in this case, the graduate engineer, unknown telephone informer, foreman, clerk of works, and resident engineer) created and used representation-bearing artefacts 'on the fly' (the work record, the teams' file, sketch and leaflet) and over different channels of communication (spoken, post, and telephone). This process was not specified in the operating procedures laid out by the construction company and was generated and interpreted creatively by the participants. In this way, the SOP functioned as an (incomplete) organising resource rather than a rigid set of instructions, and it was loosely applied in the performance of work. It did not determine the physical actions required, which were selected according to a range of social, material and spatial factors. In this respect, it was creatively interpreted, rather than followed.

What is noticeable about the example is the way that the task involves both formal (i.e., established) work practices, some of which are given in the SOP, and an ad hoc approach to collaboration. Showing that work involves formal, or explicit, and informal, or tacit, features is not itself a novel finding (e.g., Grudin 1994). Nonetheless, this activity differs significantly from the practices noted in previous DCog analyses of work, in which the formal characteristics of work are exaggerated because of the particularities of the situations examined. Recognising the unpredictability of work and the agents' use of ad hoc work practices to deal with this is a critical feature of examining activity in the analysis of loosely coupled organisational systems. This interrelationship between the formal and ad hoc practices is important in understanding activity. While Hutchins (1995a, b) and to an even greater extent Wright et al. (2000) place emphasis on the formal or proceduralised aspects of work, they do not ignore the informal aspect of collaboration entirely. However, within loosely coupled systems, the participants' ongoing orientation to the constantly shifting problem situation is central to their performance. The participants' access to a wide range of resources and the flexibility in the management of their own organisation means that agents within a functional system may exhibit many unique, situation-specific solutions to the problems they face. Application of a DCog approach in these settings must reflect this possible lack of precedents and the agents' artful use of the resources at hand.

Cycle Duration. The cycle duration on the construction and design of the temporary works was highly variable—in terms of the project as a whole, the work was expected to take 3 years; far longer than the brief cycles of navigational fix taking. However, within this project, a number of smaller problem solving systems could be identified, each of which could be examined as a functional system in its own right. These different classes of problem ranged from specific temporary works designs (such as scaffolding towers), permanent works designs (ranging from pile placement to whole bridge designs), individual concrete pours, and so on, taking place over widely differing time-scales. Even similar tasks could take place over radically differing time-scales, depending on the availability of information about the problem, the intellectual problem solving resources, and the physical resources available to the functional system. Examples of this are impossible to demonstrate with snippets of observational data, and it is perhaps unnecessary to attempt to do so. The three-year example of the project is itself an ample demonstration of cycle duration in a design and construction project; work took place over an extended period, punctuated by periods of inactivity and intensive action. Of course, not all loosely coupled systems will periods of cycle duration as long as that of the construction project described, and the duration may also vary depending on the boundary that the analyst places on the frame of the analysis. For example, investigating the entire building process might typically take several years, but to examine the initiation of plans for the building project it may be feasible to observe only for a few months, or to examine the design construction of formwork, a few weeks may suffice.

For the cognitive ethnographer, the long duration of a project presents something of an obstacle: the complete design and construction project took far longer than the time available to study it. A 3 year long project is by no means unusual in the construction industry, and similarly long project durations are relatively common in industry in general. It is therefore unlikely that a study could be made of the process as a whole within even a medium term research project. This must be contrasted with the observations and analysis of navigational fix cycle or cockpit behaviour, which could be measured in seconds or minutes. While similarities exist in the abstract nature of problem solving itself, the practicalities of this difference between the two kinds of problems and settings could not be more different. The implication of this for a DCog analysis is that the cognitive ethnography cannot be a complete study of the work process, but only a part of it. Data will have to be collected from before the period of study, and the analyst will have to envisage how it will be continued following the field study, by projecting their findings forward. In comparison to the previous studies of DCog, in many cases there may be little chance of looking at another problem cycle.

Summary. The examples presented in this study clearly demonstrate some of the differences with DCog in loosely coupled systems compared to ship navigation, aircraft cockpits and the other more limited situations in which we have seen problem solving examined through the use of DCog. Most noticeably, the work is heavily contingent, and the participants make extensive use of the wide range of representational resources that they find around them. The nature of this work was that it was highly context-dependent and unpredictable. While the engineers made plans, and organised labour and materials in an attempt to control the situation (the 'planned' component of activity), they were also constrained by the context within which the activity occurred. This involved adjusting their ongoing behaviours to the evolving circumstances on the site and making use of the resources available. This is not to say that previous studies of DCog do not account for this contingency, but that it is more exaggerated in loosely coupled organisational systems and requires a greater degree of attention.

A Comparison of Study Settings

In assessing the analytic value of DCog applied to loosely coupled systems, it will be helpful to provide a comparison of their key constituent dimensions against settings in previously published studies of DCog (see Table 2.1 for a summary). In the instance of tightly coupled systems, we illustrate this with reference to navigation as this is broadly representative of the systems to which DCog analyses have been 'applied'. It is important to recognise that this comparison is illustrative, and it is not suggested that *all* systems of one kind or another will share *all* of these characteristics. Rather, these dimensions are intended to highlight areas where there are likely to be contextual differences which impact on our ability to employ DCog effectively. Please note that these dimensions are not completely distinct and there are interdependencies between them.

Table 2.1 A comparison of key dimensions in workplace settings

Key dimensions	Tightly coupled systems	Loosely coupled systems
Access to resources	Agents and representational artefacts are restricted to a predetermined set	Agents and representational artefacts are unrestricted to a predetermined set and may change over time
Problem structure	Well-structured, identifiable and expected problems that are recurrent	A tendency toward ill-structured problems that have a high degree of uniqueness
Organisational structure and problem dynamics	Organisation has pre-specified modes of operation, characteristic of tightly constrained and managed organisations with rigid modes of operation. Division of labour is well understood and 'standard operating procedures' underpin much of normal work. Problem dynamics are relatively stable over time	Organisation's operation is only partially pre-determined; established work processes operate at an abstract level and are augmented by ad hoc approaches in interpreting these high-level operational directives. Divisions of labour are informally defined and enforced. Problem dynamics are unstable and dynamic over time
Cycle duration	Relatively short cycle for problem solving, coupled tightly to the task	Problem-solving cycle tends to be variable, with similar classes of problem taking place over widely different time scales

Access to Resources. The main difference between tightly and loosely coupled organisational systems can be most clearly seen in terms of the distinction between their access to resources. In navigation, the system is closed: the process has a fixed and restricted set of resources and external agents are not able to involve themselves in the system. In loosely coupled systems, participants may call on a larger set of resources that might not be initially specified or known to be available at the beginning of the activity.

Problem Structure. The problems that the two types of system have to solve may also be structured in different ways. In navigation the problem is ‘well-structured’ prior to its solution; the task is repetitive and agents are well practised in performing the task. In loosely coupled organisational systems, the problem is more usually ‘ill-structured’ and only becomes well-structured in performance as the agents learn about the problem during problem solving.

Organisational Structure and Problem Solving Dynamics. The methods that are used to organise the co-ordination of activity differ between tightly and loosely coupled organisational systems. In navigation, the communication pathways are (necessarily) well specified and constrained to a number of pre-defined modes of operation. These are enforced by (naval) regulations, which prescribe the division of labour on particular tasks. However, in loosely coupled organisational systems, not all of the communication pathways are well-specified prior to problem solving, and their organisation are likely to be only partially constrained by pre-determined modes of operation. While we recognise that navigational work is not always routine, this is far more exaggerated in loosely coupled systems. There may be few absolute organisational structures, and the artefacts, communication pathways and participants available are likely to change over time. Some processes may be formally specified, but many are generated in an ad hoc fashion. Formal specifications may be stipulated at a high level, but the mechanics of implementing these are frequently left to the interpretation of participants, subject to professional and legislative constraints. In loosely coupled systems, procedures can be defined at an abstract level, and it may be left to the interpretation of individuals to decide on what actions to take as circumstances change.

The changing nature of the problems faced by the navigators and by agents in loosely coupled organisational systems differs substantially, and this has implications for the way that problem solving strategies develop and enter the culture of the workplace. Navigation by triangulation is an unchanging process, developed over centuries of practice. The procedures can remain unchanged over multiple fix cycles, and although each cycle may be of short duration in itself, they are highly repetitive. In loosely coupled organisational systems, the duration of the activities that they perform is likely to be far longer—for example, contract negotiations, design development, or product testing are lengthy activities, and over time, and even within the activity, procedures are likely to develop and adapt. This is unlikely to occur *within* the fix cycle, although Hutchins shows how the development of practice and of the practitioners (1995a, p. 372) does change over time—but over multiple cycles.

Cycle Duration. The duration of the activity cycle differs substantially between tightly and loosely coupled systems. For example, in navigation, the ‘fix cycle’ is of short duration (a matter of minutes or seconds). These fix cycles are ‘snapshots’ in time, and each involves taking a bearing of the vessel’s present location. However, in loosely coupled systems, activities can occur over much longer time-spans. In one sense, this reflects the difference in the work being done, where the tasks in the navigational system are more self-contained and time-critical.

These brief descriptions suggest that work performed by tightly and loosely coupled systems may be very different, in terms of their goals, technical resources and contexts of use. However, both may be seen as information processing systems with a similar high level (cognitive) structure and it is at this level that a common approach like DCog can be applied. Nevertheless, because these systems differ in significant ways, it is likely that the methods used to examine them will also have to differ. This suggests that we cannot draw directly from previous studies of DCog and apply them directly to the analysis of other workplaces. Rather, we need to consider the differences more closely, undertaking and drawing on new studies of organisational work to help us evolve an effective analytic approach for distributed cognition into these contexts.

Reflections on the Use of DCog

The fieldwork vignettes presented above demonstrate how the computational architecture of the functional system arose through the relationships between the agents, where the transformation of problem situation into design solution involved a variety of computations. This was implemented within a socially distributed cognitive architecture that incorporated a number of actors with different skills and roles, in combination with a range of other representational media, and operating in an environment rich in resources that structured these transformations. The social and organisational co-ordinating mechanisms that brought representations together worked in concert with the physical resources and constraints of the setting to determine the outcomes of these computations.

The descriptions[1] of action in the earlier sections focus on the informational transformations that take place, as representations are re-represented in various media. In most respects this is identical to the traditional form of DCog analysis. However, in this study, snapshots of activity were collected and presented from across a very large and distributed functional system, and through time. It is not the theoretical framework of DCog that differs in this then—the analysis is still of representations and processes, transformations and co-ordinations that allow aspects

---

[1] In ethnography, this is also confusingly known as its *representation*. Use of the word ‘representation’ is distinct in meaning to its use in DCog, where it is applied to the symbol processing activities conducted in the functional system, and not by the analyst.

of the information processing structure of the functional system to be made explicit—but it does differ significantly in its practical application (limited by a small and incomplete data set). In this sense, the description is less like a task analysis, in which a single episode is examined in detail (as with traditional DCog), and is more descriptive in the traditional sense of an ethnography, where the analyst selects representative aspects of the situation. However, what is being described is how representational transformations operate within these representative samples and extend into computations across the larger functional system.

In a study of a large and complex organisational system, such as the construction example given in this paper, data collection and its analysis will need to encompass the whole range of information processing activity that the functional system is capable of performing. As researchers examining these systems, we cannot fully specify the structure of the functional system. In the terms of cognitive science, the external symbol system derived will be incompletely specified and it cannot give us the granularity that Hutchins (1995a, b) provides us with in a formal, functional structure of the activity. In practice, the wide scope of the organisational systems analysis means that the 'density' of the data collected and the coverage of the analytic methods and approach is far lower than the standard DCog approach can provide us with. In a loosely coupled activity system, we may have no means of examining all of the transformational actions undertaken. What an analysis can demonstrate is how, *in the situations observed*, the resources were applied to perform representational transformations achieving the system goal (or possibly failing to do so). Necessarily, some (possibly important) situations will not be observed (or be otherwise accessible) that are relevant. A realistic approach is therefore to do one of three things:

(i) We can reduce the activity examined to a subset of the complete activity. This is the kind of approach used successfully by DCog researchers to date. However, such an approach means that we cannot see how the activity within the wider system is performed.

(ii) We can accept that the level of granularity in the analysis will be lower, and that the descriptions of the actions performed will be less detailed. The analysis will explore the higher-level, organisational features of work rather than its practice. However, this approach means that we cannot look at the actual *use* of the representations and processes in detail. In many instances, this is what management and organisational analysts do, and in so doing, miss out the valuable roles that local practice and artefacts play in the performance of work.

(iii) We can focus on the significant actions that the participants and the ethnographer deem to be of particular importance to the performance of the functional system as a whole—whether it is particularly difficult, important to their co-ordination, where they have particular problems, or have to perform repair functions (i.e., resolve breakdowns) in particular situations.

This third option is the one that we would advise, and have used ourselves in the fieldwork described in this paper. This allows a degree of scope for the analyst to select aspects of the functional system that are particularly significant to the participants, and to do a deep analysis of the representational structures involved. However, this will not provide a complete description of actions in the functional system, and may be prone to a degree of subjective bias in the actions selected for detailed analysis. It relies to a greater extent on the importance attributed to situations by the ethnographer and participants than the standard DCog approach, but this is generally an accepted component of interpretive research (c.f. Van Maanen 1988; Hammersley and Atkinson 1995) and is not, by itself, a failing. The method of data collection—ethnography—and the analytic framework suggested provides a means of examining *aspects* of information processing. It cannot be treated as a total description of an activity, but is a means of getting a deeper understanding about that activity within its context. We further suggest several features in the analysis of loosely coupled systems that differ from the traditional approach used by Hutchins and other DCog theorists:

(1) DCog is useful in illustrating how and which resources are used, but field data can only be understood as an *instance* of problem solving activity—at another time, activities may be performed very differently because the functional system has access to a wide range of resources. It should not be presented as a complete and systematic computational description of an activity in the way that Hutchins describes navigation or cockpit activity, because of the one-off uniqueness of the activity sequences observed.

(2) Piecing together the component parts of the analysis is not a simple matter. Because of the size and complexity of the activity system being studied, the analyst may have to link a large number of interacting local functional systems (for example, the generation of the design brief or a materials order) together to form a picture of higher-level problem solving (in this case, temporary works design). When doing the analysis, the cognitive ethnographer will need to identify which local functional systems form the ‘trajectory of work’ (Strauss 1985), and to investigate these with an aim to linking them together in the final analysis of the larger functional system.

(3) Cognitive ethnography cannot always present a complete description of computational activity, but only a part of it (over time, space and tasks) that is extrapolated to cover the whole problem solving cycle (a related point to 1 and 2 above). The analyst is unlikely to be able to examine another problem cycle to see how differences occur. However, it may be possible to study another organisation or activity system to compare activities. This should not be an attempt to validate the data through triangulation, but to present a cross-cultural dataset: analysts should not attempt to produce absolute answers to questions about the organisation of activity, but to provide a *better understanding* about the social practices and organisation of the functional system in context.

(4) The nature of the problem in the situation studied may be ill-structured and, as a consequence, the participants will strive to understand the nature of this

> problem and their co-workers' orientation to the problem and each other. The focus of analysis is therefore as much on understanding ongoing system's self-organisation as it is with directly looking at problem solving on the task. The analyst will need to pay increased attention to how agents in the functional system are *made aware* of ongoing, but problem-unrelated, situation monitoring so that they can self organise. This is related to the agent's situation awareness which, we argue, must be considered as a core feature of activity, even when it is not directly associated with a specific problem-solving event. Hutchins does discuss situation awareness (1995a), but only in relation to known environmental events: for example, when landmarks come into view, when observations are shouted down the intercom, when marks are made onto tables and charts or when agents within the functional system leave or join.

Conclusion

The framework of DCog allows the analyst to examine the computational nature of an activity, and the problem solving activity that occurs over a distributed group of individuals and artefacts. The analysis lies not in the abstract processes of an activity, or in a description of the communications that made it possible, but on how the activity and its co-ordination are inter-linked through task performance. In this, we do not seek to radically alter the foundations of the DCog approach: DCog is a broad church and does not advocate adherence to a particular method or approach to gathering data about the functional system (Hutchins 1995a; Hollan et al. 2001). It is important to recognise that we are not critical of the approach used in traditional DCog; rather we show how DCog can be applied to situations in which the traditional DCog analytic *approach* (exemplified by Hutchins 1995a) would not be possible or appropriate.

In some ways, what we are proposing in terms of the analysis of open and loosely coupled systems is not dissimilar to that commonly performed in ethnomethodological studies of action (Garfinkel 1967; Heritage 1984). The scale of the analytic frame, the form of the data collected, and the ways that actors and artefacts may be enrolled into systems in an ad hoc way can be seen as similar. Yet these approaches are quite distinct as theoretical approaches and in what they offer in an analysis. The idea that a system might be 'cognitive' in any way is something of an anathema to ethnomethodologists (Heritage, ibid.), who take a distinct different perspective to cognitive science (indeed, they are explicitly atheoretical in their approaches to examining human action) and hold well-articulated arguments against cognitivism and the idea of goal-directed approaches to action.

A core difference between the sorts of distributed cognitive analysis performed in this study and the sort of study performed in tightly coupled systems is that we do not attempt to describe a complete computational structure for the functional system. Rather, we look to the framework of DCog as a means of exposing the more highly relevant information processing aspects of the activity (as judged by the fieldworker or workplace analyst). This is in contrast to much of the existing work in DCog, which attempts to produce well-defined accounts of functional systems that capture most, if not all, of the information processing characteristics inherent in them.

While the form of analysis advocated here for loosely coupled systems cannot be as precise as that produced by an analyst working with more tightly coupled systems, the approach is clearly of relevance to organisational theorists and systems designers. It changes the remit of the analyst from providing a near complete specification of a problem solving activity into one that demonstrates how to focus the analysis of ethnographic data onto the informational component of an activity. For applied researchers, this focus on information processing to show how representations and processes can be appropriated for use in performing and co-ordinating distributed problem solving activity has direct implications for systems design: it foregrounds the role of information and unites the social, organisational, psychological, situational, and artefactual elements that contribute to problem solving. The value of this is that analysts are able to see the interconnectedness of these factors in the activity, and to envisage how any interventions might have beneficial or harmful impacts on the activity system's information processing characteristics.

Acknowledgements The article has greatly benefited from discussions with Jon Hindmarsh, Kenton O'Hara and Barry Brown.

References

Alač, M., & Hutchins, E. (2004). I see what you are saying: Action as cognition in fMRI brain mapping practice. *Journal of Cognition and Culture, 4,* 629–661.

Artman, H., & Wærn, Y. (1999). Distributed cognition in an emergency co-ordination center. *Cognition, Technology & Work, 1,* 237–246.

Blanford, A., & Furniss, D. (2006). DiCoT: A methodology for applying distributed cognition to the design of teamworking systems. In S. Gilroy & M. Harrison (Eds.) *DSV-IS'05 (International Workshop on Design, Specification and Verification of Interactive Systems), LNCS 3941* (pp. 26–38). Springer-Verlag.

Engestrom, Y., & Middleton, D. (1996). *Cognition and communication at work: Distributed cognition in the workplace*. Cambridge: Cambridge University Press.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs NJ: Prentice-Hall.

Giere, R. N. (2004). The problem of agency in scientific distributed cognitive systems. *Journal of Cognition and Culture, 4,* 759–774.

Grudin, J. (1994). Eight challenges for developers. *Communications of the ACM, 37,* 93–104.

Halverson, C. A. (1995). *Inside the cognitive workplace: New technology and air traffic control.* Ph.D. Thesis. Department of Cognitive Science, University of California, San Diego.

Halverson, C. A. (2002). Activity theory and distributed cognition: or what does CSCW need to do with theories? *Journal of Computer Supported Cooperative Work, 11,* 243–267.

Hammersley, M., & Atkinson, P. (1995). *Ethnography: Principles in practice* (2nd ed.). London: Routledge.

Hazelhurst, B. L. (1994). *Fishing for cognition: An ethnography of fishing practice in a community on the west coast of Sweden*. Ph.D. Thesis. Department of Cognitive Science, University of California, San Diego.

Heath, C., & Luff, P. (1991). Collaborative activity and technological design: task coordination in London Underground control rooms. In *Proceedings of the 2nd European Conference on Computer Supported Cooperative Work* (pp. 65–80), Amsterdam, The Netherlands.

Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge and New York: Polity Press.

Holder, B. E. (1999). *Cognition in flight: Understanding cockpits as cognitive systems.* Dissertation, Cognitive Science, University of California, San Diego.

Hollan, J. D., Hutchins, E., & Kirsh, D. (2001). Distributed cognition: A new foundation for human-computer interaction. *ACM Transaction on Computer-Human Interaction, 7,* 174–196.

Hughes, J. A., Randall, D., & Shapiro, D. (1992). Faltering from ethnography to design. CSCW 92: Sharing perspectives. *Proceedings of the Conference on Computer supported Cooperative Work* (pp. 115–122). N.Y.: ACM Press.

Hutchins, E. (1991). Organizing work by adaptation. *Organization Science, 2,* 14–39.

Hutchins, E., & Klausen, T. (1991). Distributed cognition in an airline cockpit. In Y. Engeström & D. Middleton (Eds.), *Communication and cognition at work* (pp. 15–34). Cambridge: CUP.

Hutchins, E. (1995a). *Cognition in the wild*. Bradford: MIT Press.

Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science, 19,* 265–288.

McGarry, B. (2005). *Things to think with: Understanding interactions with artefacts in engineering design*. Ph.D. Dissertation, University of Queensland, St Lucia, Brisbane, Australia.

Murray, D. (1993). An ethnographic study of graphic designers. *Proceedings of the third European Conference on Computer-Supported Cooperative Work* (pp. 295–309). Netherlands: Kluwer.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs: Prentice-Hall.

Norman, D. A. (1993). *Things that make us smart*. Reading, MA: Addison-Wesley.

Perry, M. (2003). Distributed cognition. In J. M. Carroll (Ed.), *HCI models, theories, and frameworks: Toward an interdisciplinary science* (pp. 193–223). San Francisco: Morgan Kaufmann Publishers.

Resnick, L. B., Levine, J. M., & Teasley, S. D. (1991). *Perspectives on socially shared cognition*. Hyattsville, MD: American Psychological Association.

Rogers, Y., & Ellis, J. (1994). Distributed cognition: An alternative framework for analysing and explaining collaborative working. *Journal of Information Technology, 9,* 119–128.

Salomon, G. (1993). No distribution without individual's cognition: A dynamic interactional view. In G. Salomon (Ed.), *Distributed cognitions* (pp. 111–138). Cambridge, Mass.: Cambridge University Press.

Simon, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence, 4,* 181–204.

Simon, H. A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge, Mass.: MIT Press.

Strauss, A. (1985). Work and the division of labour. *The Sociological Quarterly, 26,* 1–19.

Van Maanen, J. (1988). *Tales of the field: On writing ethnography*. Chicago: University of Chicago Press.

Wright, P. C., Fields, R. E., & Richardson, M. D. (2000). Analyzing human-computer interaction as distributed cognition: The resources model. *Human Computer Interaction, 15,* 1–41.

Chapter 3
Distributed Cognition at the Crime Scene

Chris Baber

Abstract The examination of a scene of crime provides both an interesting case study and analogy for consideration of Distributed Cognition. In this paper, Distribution is defined by the number of agents involved in the criminal justice process, and in terms of the relationship between a Crime Scene Examiner and the environment being searched.

The examination of a crime scene is subject to all manner of legal, ethical and scientific imperatives, and the evidence collected will be subjected to inspection by a variety of individuals with different intentions, skills and knowledge. In this paper, I will suggest that Crime Scene Examination presents an interesting and challenging domain in which to consider the notion of Distributed Cognition for the simple reason that it is not always apparent where the act of 'cognition' is situated. The ultimate aim of the criminal justice process, of course, is to acquire evidence which can be combined with information from other sources in order to produce a case that can be tried in Court. Contrary to its representation in popular fiction, the examination of a crime scene is unlikely to yield evidence that immediately links a suspect to a crime. Rather, the collection of evidence is part of a complex web of investigation that involves many individuals, each considering different forms of information in different ways. Thus, the paper begins with a cursory description of the role of the Crime Scene Examiner (CSE) within the criminal justice process.

The CSE is part of a much larger investigative system, each member of which has their own skills and roles (Smith et al. 2008). In a sense, Crime Scene Investigation involves sets of ad hoc teams pursuing independent goals with quite limited overlap (Smith et al. 2008). Thus, there is typically a demarcation between roles. Having said this, the nature of this demarcation has been subject to significant shifting over the years, with the ongoing digitisation of Crime Scene Examination leading to further changes. For example, there used to be a specific role of Crime

---

C. Baber (✉)
Electronic, Electrical and Computer Engineering, The University of Birmingham,
Birmingham B15 2TT, England
e-mail: c.baber@bham.ac.uk

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_3

Scene Photographer whose function was to capture and process images of the crime scene (either prior to evidence recovery or at stages during the recovery process, depending on the nature of the crime). However, with the growing use of digital cameras by CSEs, this role has (in some Police Forces) changed. This has the interesting implication that the function of a photograph taken by the Crime Scene Photographer was to capture the scene as clearly as possible in order to aid discussion of the scene in Court (or during subsequent investigation), but the function of a photograph taken by the CSE *could* be to illustrate the evidence recovery process; I suggest this because the capturing of images by the CSE is *part* of the activity being undertaken rather than the sole focus of the activity. Whether or not similar changes might arise in terms of specialised analysis of fingerprints, footwear marks, DNA and other evidence is a matter of continued debate. For the time being, these analyses are generally performed by Forensic scientists rather than by CSEs. This means that one of the primary roles of the CSE is the recovery of evidence and its transportation in a usable state to the laboratory of the Forensic scientist. How this recovery and transportation is performed, and how closely the Forensic scientist and CSE cooperate depends very much on the nature of the crime being examined. For much of our work, we have focused on what is called ‘Volume Crime’ (e.g., robbery, burglary), as opposed to ‘Serious Crime’ (e.g., murder, rape, kidnapping). In Volume Crime, it is likely that the recovered evidence is passed on to the Forensic Scientist via a third party (sometimes called the ‘Evidence Manager’). This means that any information pertaining to that item needs to be carefully and comprehensively recorded by the CSE prior to depositing with the Evidence Manager. It is this combined process of recovery, storing, labelling and transportation of evidence that forms the basis of several forms of computer-based CSE support (i.e., evidence management systems). Before exploring this further, we consider the archetypal detective and his approach to investigating crimes.

Sherlock Holmes and Reasoning About Crime

Sherlock Holmes tells a visiting stranger “You have come up from the South-West I see” observing that the “…clay and chalk mixture which I see upon your toes caps is quite distinctive.” (Conan Doyle 1989, p. 176, *The five orange pips*). This ability to draw correct conclusions from visual evidence is one of the hallmarks of Holmes’s powers, and implies a particular form of reasoning. Holmes’s method is a form of *induction* which involves the careful observation of the environment in order to develop hypotheses and then performing a process of elimination among a number of alternative possibilities, that is, “…eliminate all other factors, and what remains must be the truth.” (Conan Doyle 1989, p. 66, *The sign of four*). So that, “one simply knocks out all the central inferences and presents one’s audience with the starting-point and the conclusion, [so that] one may produce a startling, though possibly a meretricious, effect.” (Conan Doyle 1989, p. 583, *The adventure of the dancing men*). He would often present his conclusions as the result of deduction (i.e., 'Elementary, my dear Watson') and imply that he was able to draw a conclusion from general principles to a specific observation; indeed, Holmes would often refer to his method as *deduction*. One could argue that Holmes was attempting to apply a deductive method (through his exposition of premises) but was hampered by Conan Doyle's insistence of continuing to add extra pieces of evidence, which forced him into an inductive method.

This distinction between induction and deduction is based on a broad characterization of the approaches as rival positions, namely induction as 'observations leading to theory', and deduction as 'theory guiding observation'. In reality it can be difficult to separate the two, and difficult to conceive of the 'pure' application of induction (which would involve the compiling of observations in a manner which was theoretically agnostic, and the subsequent development of a theory which was *solely* based on those observations). One would assume that observations will be, in some sense, selective and that this selectivity could be tuned by attention to specific aspects of the environment. The point of this discussion is to raise a key issue for Crime Scene Examination; there is a supposition that the work of the CSE involves the 'harvesting' of materials which would then be analysed by Forensic Scientists. CSEs are supposed to maintain neutrality in terms of collecting evidence and to conduct their work in an inductive manner, because any sense in which they are interpreting the scene could be construed as a potential for bias in the investigation. Of course, Holmes never had to face such accusations because, as a literary character, he was not guilty of bias (only of revealing the information given to him by his author) and did not have to justify his interpretations under cross-examination in Court. The question of how Crime Scene Examination treads the line between induction and deduction is explored later in this paper; before this we will consider the notions of Distributed Cognition that underlie our studies.

Distributed Cognition

The notion that cognition can be 'distributed' has been developed over the past couple of decades (Artman and Waern 1999; Artman and Garbis 1998; Busby 2001; Flor and Hutchins 1991; Furness and Blandford 2006; Hollan et al. 2002; Hutchins 1995a, b; Hutchins and Klausen 1998; Perry 2003). While I suggest that Crime Scene Examination necessarily involves several agents performing cognitive activity, this is not to argue that this results in an 'extended mind' across these agents; as Dror and Harnand (2008) point out, to argue for an extended mind is analogous to arguing for extended migraine—just because an event occurs in one brain does not inevitably mean that other brains will share this event. Dror and Harnand's (2008) argument is that one should not separate cognitive states from mental states. This criticism raises a core problem for the notion of 'Distributed Cognition', because it implies that cognition cannot be 'distributed' across agents because one cannot share mental states. A primary assumption of 'distributed cognition' is that it is not 'cognition' which is distributed so much as objects-in-the-world, which play a role in supporting, structuring and aiding the activities of cognition. “A main point of departure from the traditional cognitive science framework is that, at the ‘work setting’ level of analysis, the distributed cognition approach aims to show how *intelligent processes in human activity transcend the boundaries of the individual actor*. Hence, instead of focusing on human activity in terms of processes acting upon representations inside an individual actor’s heads the method seeks to apply the same cognitive concepts, but this time, to the interactions among a number of human actors and technological devices for a given activity.” (Rogers 1997, p. 2). This quotation hints at two notions of an ‘extended mind’. For example, some theorists claim that the mind can become ‘extended’ through its interactions with the environment, for example “…certain forms of human cognizing include inextricable tangles of feedback, feed-forward and feed-around loops; loops that promiscuously criss-cross the boundaries of brain, body and world.” (Clark 2008, p. xxviii). Thus, as we shall in the section entitled ‘Inspection and Examination’, objects-in-the-world (and the representations made of them) form resources-for-action through their ability to afford specific responses. In addition, the crime scene examination process also features a distribution of tasks. What is particularly interesting, from the point of view of Distributed Cognition, is that the process of ‘find-recover-analyse-interpret-conclude’ is divided between two or more people, with quite limited communication between them. The CSE might perform the ‘find-recover’ tasks to gather potential evidence and then submit this for the ‘analyse-interpret’ tasks by a Forensic Scientist, who would then pass the results on to the Officer in Charge of the case with a probability score to guide the preliminary ‘conclude’ tasks. The Officer in Charge would then combine this evidence with other information to raise an hypothesis and add this to a Case file which would be passed to the Crown Prosecution Service. This hypothesis, if maintained, would then be tested in Court by Barristers presenting a case for and against an individual.[1] Each step of this process would be documented and conclusions drawn in such a way as to avoid potential bias.

One could draw an analogy between ‘extended mind’ and the debate over ‘broad’ and ‘narrow’ mental content in Philosophy. The notion of ‘narrow’ content might assume that a person’s belief about something could be defined entirely by their intrinsic characteristics (and would not change with any changes in their environment). The notion of ‘broad’ content, on the other hand, is inextricably tied to the person’s environment. For example, Putnam (1975) contrasted beliefs about the concept ‘water’ between Earth and ‘Twin Earth’. Twin Earth was exactly the same as Earth, with the exception that the chemical properties of that element termed ‘water’ were different (although the observable properties were the same on Earth and Twin Earth). Putnam’s (1975) claim was that, given identical individuals on Earth and Twin Earth, when either spoke about ‘water’ they would be referring to something different. This means that the intrinsic characteristics of these two identical individuals would not be sufficient to determine the meaning of the word

---

[1] This example follows the legal system in England and Wales; while other countries will follow different processes, the point is that several people are involved in the interpretation of evidence.

'water', but that there needs to be some reference to external environment. This leads Putnam (1975) to make the well-known assertion that "…meanings' just ain't in the head" (p. 227).

Relating this discussion to the earlier contrast between Sherlock Holmes and contemporary CSE, we could suggest that Holmes represents the application of 'narrow' content; the world and its machinations exist solely through his (or rather, Conan Doyle's) description of them and this description cannot be challenged (simply because the stories rarely include the opportunity to develop alternative explanations). In contrast, the CSE is involved in the application of 'broad' content; the world is represented as evidence which is passed between different people who can offer different interpretations to bear on it. From this perspective, the question becomes a matter of how representations are used rather than a matter of *individual* interpretation (because these interpretations will always, in an adversarial legal system, be open to dispute).

Distributing Examination

While Sherlock Holmes provides an entertaining version of logical analysis (and serves as a template for contemporary television equivalents), his approach has many differences with modern Crime Scene and Forensic Examination. Obviously, Crime Scene Examiners do not have the benefit of the omniscient author guiding the discovery and interpretation of evidence, nor do they have the opportunity to present their findings to an informal (usually incredulous) gathering of people, as could Holmes. More importantly, Holmes's form of inductive reasoning requires the probabilistic elimination of competing hypotheses to explain a well-defined piece of evidence. The notion of a well-defined piece of evidence concerns the relationship between recognizing something as having potential evidential value and the interpretation of that evidence in terms of other information. For Holmes (and his modern, fictional counterparts) this all takes place in the head of one person; so the processes are typically assumed to involve the mental states of a single individual.

Crime Scene Examination can be considered 'distributed', in a trivial sense, in that several people are involved in the interpretation of evidence, each providing a particular perspective on this interpretation. What we see in Sherlock Holmes is a literary representation of the many-headed being of the criminal justice process in the body of a single individual. As crime scene examination grew increasingly 'scientific' so the division of tasks into discrete specialisms (each with a defined skill set) developed (Horswell 2004). Thus, it is typical for the Crime Scene Examiner and Forensic Scientist to have followed different career paths and have different skill sets (and, furthermore, for there to be a growing variety of specialisms within Forensic Science). Two further factors in the 'distribution' of Crime Scene Examination arise from the 'civilianisation' of CSE activity (the recruitment of personnel to this function from outside the Police Force) and the establishment of specific CSE units (outside the operation of separate Police stations). Each of these factors can be related to imperatives of economic and efficiency gains, but they have a bearing on how knowledge of criminal behaviour is shared and applied. For example, an understanding of criminal behaviour, gained over years of policing, could help interpret evidence; but recruiting civilian staff to these posts might remove the opportunity to gain knowledge and experience from policing. This could be dealt with through the training and exposure of new CSE personnel, or through the integration of CSE activity with other police activity. This relates to the second point, namely the removal of a CSE from local police stations to centralised services, implies the need for a means of sharing experiences and knowledge. Thus, if there is a set of similar cases in an area (say a string of burglaries with similar ways of gaining access to a building), then one would expect a link to be made between them. However, if each case is investigated by different individuals, then it might not always be possible to explore such links.

What is happening in Crime Scene Examination is the mediation of cognition through the collection, manipulation and dissemination of a variety of artifacts; each artifact is interpreted in particular ways by the agents who come into contact with it. My argument will be that, for the various agents involved in this evidence chain, each artifact can ‘afford’ a particular set of responses, that is, the artifacts are resources for action, and the actions will be recognized by different agents according to their training and experience. I am using the notion of ‘afford’ in the sense introduced by Gibson (1977, 1979), as a form of perception-action coupling in which the physical appearance of an object in the world supports particular physical responses (e.g., a pebble ‘affords’ grasping in the hand). Thus, the design of artefacts that are used in a work environment become changed by their use, and these changes provide cues for subsequent use (Bang and Timpka 2003; Nemeth 2003; Seagull et al. 2003). What makes this a challenging domain for discussing Distributed Cognition is that the manipulation of an artifact by one agent might have a significant bearing on the state of the artifact, which could interfere with the activity of other agents, e.g., a simple example would be the need to preserve a crime scene so as to protect evidence from contamination conflicting with the need to retrieve specific items of evidence, or the need to dust a surface to reveal fingermarks conflicting with the need to photograph the scene.

Inspection and Expectations

In their study of Crime Scene Examination, Schraagen and Leijenhorst (2001) recorded verbal protocols of the examination of a staged crime scene. They suggested, for the analysis of these protocols, that the experienced Crime Scene Examiner develops a narrative of the crime, for example considering how a person might have gained access to the building, what path they might have followed, what actions they might have performed etc. This narrative would probably be intertwined with the search activity, such that the narrative would influence the search and the search would influence the narrative. In a similar vein, Ormerod et al. (2008) suggest that "…expert investigators … [call] … upon internalized cognitive frames relating to human behaviour that allow them to generate expectations about the actions and responses of others in real time" (Ormerod et al. 2008, p. 82).

In studies using ASL MobileEye, a head-mounted eye-tracking system, we asked Crime Scene Examiners to inspect a set of staged crime scene. In one study we compared performance of three experienced Crime Scene Examiners and three Undergraduate students to search the same room under the same conditions. Of the many obvious and striking differences between the two sets of recordings, we noted that the students had a tendency to search only around locations that they believed to have links with stolen items—and so their narrative was focused solely on the loss of objects. The Crime Scene Examiners had a far more detailed narrative to guide their search and, as the stills from one recording shown below illustrate, spent a substantial part of their time looking at the door and noting possible evidence that could be recovered, for example blood stains near the latch, tool marks made by a chisel on the door frame, a footprint on the outside of the door.

Discussion with the Crime Scene Examiners showed how experience played a key role in deciding where to look for evidence and how best to examine the scene. For volume crime, the Crime Scene Examiner might walk the scene with the victim in the first instance, and then return to key locations to look for possible evidence. There was some debate as to what should be the first location to search. Standard practice might say that one begins with the Point of Entry and examines that thoroughly. In Fig. 3.1, the Point of Entry involved forcing an office door, possibly with a tool that had a sharp end, such as a chisel, which resulted in cuts around the latch. Fingermarks on the door could have been left during entry (or exit) and suggest that the entrant had cut the right thumb. Comparison between experienced

Fig. 3.1 Stills taken from mobile eye-tracker worn by Crime Scene Examiner inspecting a staged break-in (fixation indicated by cross in *thick circle*)

CSEs and the untrained Engineering students with *no* experience of CSE work showed clear distinctions in search pattern; whereas the students all walked into the room without looking at the door, the CSEs all spent around 20% of their total search time inspecting the door before proceeding to the rest of the room. There are two plausible explanations for this. The first is that this scene (which had been staged to replicate an office break-in) had conspicuous evidence on and around the door. However, this evidence was not so conspicuous that the students noticed it. The second is that the CSEs expect to find evidence at Point of Entry and so attend to this in detail. The CSEs, after the study, stated that this approach was 'intuitive' and 'just felt right'. In their discussion of intuition in problem solving, Dreyfus and Dreyfus (1986) noted that "intuition is the product of deep situational involvement and recognition of similarity…; [and becomes expertise when] not only situations but also associated decisions are intuitively understood." (Dreyfus and Dreyfus 1986, p. 18). This notion is analogous to Klein's notion of Recognition-Primed Decision-making (Klein et al. 1986). In Recognition-Primed Decision-making (RPD), one can infer three broad approaches that the decision-maker might follow; (i) the situation is recognized as 'typical' and an associated set of activities would be brought to mind; (ii) the situation is defined in terms of core features, each of which would be developed in terms of (i); and (iii) the situation is unusual, and the person might mentally explore alternative strategies prior to committing to a set of activities. This study, and discussion with the Crime Scene Examiners, implies that the situation was defined in terms of (ii), and that each aspect would be considered in terms of a set of activities. The Point of Entry was explored in terms of recoverable DNA, fingermarks, and toolmarks (possibly in this order because each might be considered to have different levels of permanence and need to be recovered quickly). In a similar manner, Flin et al. (2007) have suggested that operational policing involves recognition of situations and the subsequent elicitation of appropriate response scripts, so this example of CSE suggests a three-step process by which a set of 'typical situations', such as Point of Entry, are used to guide search of a scene, which then leads to attention to items of potential evidential value, and then interpretation of these items. Thus, we could reverse Klein's RPD to describe the activity of the CSE as Decision-Primed Recognition. This is not a huge step in terms of Klein's notion of RPD because it simply follows the perception-action cycle that RPD implies: The recognition of features in the environment are responded to in terms of decisions based on previous experience, and these decision, in turn, can help shape expectations of what to look for in the environment (and to help interpret what one is looking at).

A second study concerned compared first students on a crime scene examination and forensics degree and experienced crime scene examiners (Baber and Butler 2012). In one condition, there was a search of a ransacked office (again the scene was staged). Figure 3.2 shows a set of stills taken from an experienced Crime Scene Examiner opening the office door and immediately noticing a black mark on the floor (a), closer inspection indicates that this is a footwear mark (b) and, during the course of subsequent searching a plastic bag is found under a table and a pair of shoes found in the bag—the shoes have a black substance on their sole and the tread

Fig. 3.2 Series of images from eye-tracking worn by experienced CSE inspecting a ransacked office

looks similar to that in the footwear mark (c). The scene had been staged to look as if an opportunistic thief had broken into the office and stolen money from a petty-cash tin (which was left open on top of the desk). However, in a twist in the scenario, we had staged the scene to actually reflect an 'insurance job', that is, the office's owner had staged the crime to claim on his insurance for loss of cash, personal possessions and some computing equipment.

Most of the evidence in the scene could have been used to support the conclusion of an opportunistic crime, which was the conclusion of all 5 students and 2 of the CSEs. There were three crucial pieces of evidence which pointed to the alternative conclusion (the shoes, as shown in Fig. 3.2; the fact that the window looked to have been forced but with no obvious evidence of it being used as a point of exit, particularly as it was some 15′ off the ground; the order in which the desk drawers had been opened[2]).

The stills in Fig. 3.2 show an additional aspect of the CSEs exploration of the scene. As well as being guided by their experience of likely places to search for evidence, they need to maintain a running commentary of recovered evidence so as to be able to compare subsequent finds. Interestingly, the two CSEs who did not link the shoes to the footwear mark had previously dismissed the marks as 'smudged' and 'not worth recovering'. This implies that the mark was no longer part of their running commentary and so the potential value of the shoes was not explored. The question of how a 'running commentary' is developed and indexed during a search activity could be worth further investigation. Studies of Distributed Cognition demonstrate ways in which, objects-in-the-world structure cognition. Often these objects-in-the-world are purpose-built to support specific cognitive activities, or are adapted from existing objects. Researchers would then either focus on the design of such objects, and their ability to support cognition or at ways in which activities result in the modification of objects. Crime Scene Examination represents a special case, in that the objects-in-the-world to which the person attends have been neither designed nor adapted to suit a specific cognitive activity.

---

[2] In order to prevent one drawer obscuring the contents of the next, and in order to prevent the need to close drawers, the experienced criminal is likely to open drawers from the bottom up—but in this scene, we had obviously opened them top down.

Rather, the objects have to be discovered by the person and then interpreted in terms of their relevance to the task of gathering evidence. In this manner, the tasks of discovering objects-in-the-world that could have evidential value can be considered a form of recognition-primed decision-making.

Evidence Recovery

As mentioned previously, one requirement of Crime Scene Examination is to select items that *could* be of evidential value. This means not only finding visible items, but also preparing surfaces so that less visible, or latent, items can be revealed. Figure 3.3, for instance, shows how a surface can be prepared to lift fingerprints. In this instance, the item being inspected (a glass bottle) is being dusted with aluminium powder using a brush. The brush is applied to the item using a swirling motion to ensure a light, even coverage. The process involved a period of brushing (for around 10 s), followed by a visual check (for about 5 s in which the bottle was gently rotated to catch light falling on any revealed marks), and then a repeated period of brushing prior to the use of tape to lift the revealed marks (or, more recently, the use of high-resolution digital photography to capture the marks) to transport them to the laboratory. In some instances, the visual check might be supplemented through the use of a handtorch which is shone orthogonally to the powdered surface. In the inspection shown in Fig. 3.3, the torch was not used but the CSE could be seen to be rotating the bottle to catch available light during the visual check phase. Concurrent verbal protocol during the search suggested that the CSE initially concentrated on two areas that were anticipated to reveal marks—and there was an assumption that each area would reveal different types of mark. Around the neck of the bottle, the search was initially for marks from fingertips and thumb holding the bottle vertically (as if carrying it) and around the middle of the bottle the search was for marks of the bottle resting across the middle of the fingers and being controlled by the thumb. Thus, a schema of how the bottle could have been used influenced the initial search.

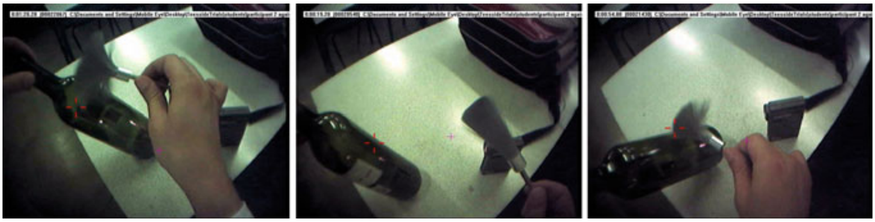


Fig. 3.3 Dusting for fingermarks

While there are procedures in place for the recovery and analysis of finger marks, work by Dror et al. (2005) highlights how their interpretation could be biased with the provision of additional contextual information. In this study, contextual factors were manipulated by the story and photographs that were used to explain the source of the fingerprints, for example crimes with no physical harm to the person versus crimes with extreme physical harm. The study showed that in cases where the fingerprints were unambiguously different, there was little effect of context. When the fingerprints were ambiguous, namely when the certainty as to whether they were the same of different decreased, then the contextual factors seemed to play a role in increasing the likelihood of seeing a match. However, this effect was only observed for the context in which extreme physical harm featured in the background story. The study suggests that in cases where there might be some uncertainty as to whether fingerprints match and where the crime is extreme, that matching might be influenced by context. This also suggests that, while the use of a narrative to guide the collection of evidence might be beneficial, it can also bias interpretation and, by implication, search. This raises the potential (and, perhaps, often unexplored) question of how recognition-primed decisions can become biasing rather than supporting, particularly in terms of expectancy bias. This also highlights the importance of maintaining as neutral a description in crime scene reports associated with recovered evidence as possible, and shows why the inductive approach is preferable for the CSE; even if the final ‘theory’ to which the evidence leads is not developed by the CSE but by other people in the criminal justice process.

Evidence Sharing

The preceding discussion implies that the search of a scene is guided by experience, expectation and the ability to recognize items of evidential value. In this respect, the notion of Distributed Cognition can be interpreted in terms of the use of objects in the world as resources-for-action. The Crime Scene Examiner recognizes objects as resources-for-action which may well differ from untrained observers. For example, while the untrained observer might assume that a pane of glass in a window could yield fingermarks, they might be less inclined to immediately assume that it could also yield footwear marks, and still less inclined to recognize its potential for yielding DNA (the latter two could arise from someone climbing in through the window, or from pressing their forehead against the window to see if anyone is at home).

So far, this description looks very much like a process that involves the mental states of an individual; the CSE interprets the scene, recognizing objects as resources-for-action, and then recovers the evidence. However, what makes the Crime Scene Examination process different from a Sherlock Holmes story is that the CSE submits the evidence for interpretation by other people. Indeed, it is unlikely for the CSE’s notes and reports from the scene to include any deduction.

Rather the report will be as descriptive as possible. This representation, of the scene and its evidence, is passed along the recovery train. So we have a set of processes that could ostensibly represent the stimulus (or input) to a cognitive processing system. This processing is (formally) undertaken by people other than the CSE.

Once evidence has been recovered, it is placed in appropriate bags (or containers), labelled and passed on the Forensic Laboratory for further analysis. This step in the process requires some means of maintaining accurate records of who has handled the evidence, as well as the accumulation of the results of analyses. This relates to a point made earlier, that the 'distributed' nature of the Crime Scene Examination process can make this process somewhat disjointed, in that it is not uncommon for the Forensic Scientist in the laboratory to have very little information on the item recovered. One could make a strong argument that this lack of information helps an analysis to be as objective as possible, by focussing only on the item at hand (and avoiding the potential for bias that Dror et al. (2005) demonstrated). On the other hand, it might be useful to have some knowledge of the item in situ, so as to decide how best to conduct analysis. If the Forensic Scientist had recovered the item herself then such information would be recalled by her, but when it is delivered in a batch of bags then such information is not obviously available. As an example of why this could be problematic, consider a finger-mark left on a window. This mark might not be detailed enough to form a print, but could indicate whether the window has been forced up or whether someone climbed down the window—knowing the orientation of the mark on the window can help decide how best to analyse it, but this might not have been provided in the evidence log.

Reporting and Disclosure

In previous discussions of Crime Scene Examination, Baber et al. (2006a, b) consider the manner in which narratives are passed through the evidence chain. The argument was that different people in the evidence chain develop narratives (both formal and informal) that summarise the key aspects of their interpretation of the events and environment. Thus, a victim or witness might provide an account of the events as they recall; although, of course, the nature of eye-witness testimony is notoriously contradictory and prone to error (Wells and Olson 2003). Each account would develop a particular narrative, emphasizing the aspects that the witness feels was relevant, and attempt to maintain an internal coherence and consistency (but which might differ from other accounts). Interviewing of suspects, in part, involves comparing different narratives (from the suspect versus a synthesis of the witness statements which maintains coherence and consistency). In this context, the role of forensic evidence becomes merely a tool to resolve any ambiguities in these accounts. However, of course, forensic evidence has become increasingly significant in investigations (to the extent that it is often given priority over narratives because of its assumed objectivity in comparison with the obvious subjectivity and potential for bias in the narratives). We propose that each step in the criminal justice process involves the production of narrative. There are the formal narratives that are structured by the reporting procedures and forms that are used to record investigations and analyses. This would lead to a set of reports, from Crime Scene Examiners and Forensic Scientists, which are written in a scientific style and which record details in as objective a manner as possible. Such narratives would then be subjected to scrutiny in Court in terms of the methods used to perform the analysis and the interpretation of the results. On the other hand, there are informal narratives that are passed on through discussion with agents involved in the investigation (say, between an attending officer and a victim, or between the attending officer and the crime scene examiner). These tend not to be recorded for several reasons. First, as discussed below, Laws of Disclosure mean that anything which has a bearing on the case needs to be available to both Defence and Prosecution so as to maintain fairness and balance. Second, and perhaps more importantly, much of this informal narrative could be said to involve the development of formal narrative, for example, an experienced attending officer might speak with a victim to calm or reassure them prior to taking a formal statement, and during this process the victim might have several partial accounts of what has happened but be seeking to reconcile this into a single.

The final decision of the relevance of an item of evidence is made in Court during the hearing. However, an initial assessment will be made (in the UK) by the Crown Prosecution Service which will evaluate the evidence that is being presented in support of a case and decide whether it is suitable. This raises one of the key dilemmas in evidence recovery and relates to the Laws of Disclosure. Basically, these Laws of Disclosure state that anything that has been collected as part of the investigation can be made available to both Prosecution and Defence (even if it is not presented at Court). This raises two issues for this discussion. First, the adversarial nature of the Justice System (in the UK and many other countries) means that the ‘distributed cognition’ involves not only cooperation and collaboration (in terms of several people contributing to a common goal) but also conflict (in terms of two parties attempting to prevent each other from achieving their goal). I am not sure that there are many other areas of distributed cognition research which come up against this problem (although, of course, one can imagine many examples from military and law enforcement). Second, the process often involves a number of different forms of analysis and interpretation. In Baber et al. (2006a, b) we referred to these forms as formal and informal narratives, and suggested that there was a continual development of narratives, along several lines, over the course of an investigation and that very often these narratives might not connect.

Conclusions

In this paper, I suggest that, for Crime Scene Examination, cognition is distributed in three senses. First, there is the distribution of attention between the activities involved in searching, recovering and reporting. Second, there is the distribution of cognition between CSE personnel and the scene itself; the manner in which the scene is examined provides hints and cues to what evidence to recover, and interrupting this process (through the need to complete lengthy reports) could disrupt this process. For this activity, the environment and objects it contains, become resource-for-action that the experience and training of Crime Scene Examiners allow them to interpret in ways which might be different to that of the untrained observer. Furthermore, the manner in which recovered items are passed from one person to the next in the evidence chain can modify the role of these items as resources-for-action; each step in the process interprets the information from the previous step in terms of additional knowledge and information. Third, there is the distribution of information between CSE personnel and other people involved in the investigation. The notion of formal and informal narrative, and their development through the criminal justice process, sees these narratives as additional resources-for-action.

A 'weak' view of the Distributed Cognition argument might claim that what is being distributed is the collection of objects upon which the act of cognition can be focused. This would require objects-in-the-world to play a fairly passive role in the process of cognition and for them to function as vehicles for the storage or representationof information. The artefacts allow users to off-load information and also a record of previous activity. In this version, the objects have their states altered by the actions that their users perform on them (e.g., through note-taking, folding or other markings). Furthermore, not only do these objects provide a means of recording and storing information, but their design affords (or influences) the actions of the person using them.

A 'strong' view of Distributed Cognition, posits that it is the tasks involved in cognition which are being distributed. One way in which the activity of the CSE differs from some of these domains, is in the initial definition of objects-in-the-world, and for these objects to be 'revealed' in order to be recovered. This would regard the role of the CSE is primarily one of induction, or rather, as one of providing the set of alternatives upon which a process of induction could be applied. I would suggest that the act of induction takes place in the Court (or at least in the Crown Prosecution Service which decides whether a Case can be presented to Court). Prior to this act of induction, there are initial acts of deduction which are formally assigned to the Forensic Scientists, in their analysis and interpretation of evidence, but also informally applied by the CSE in the decision as to where to look and what to recover. In this view, one would expect agents and objects-in-the-world to be more active and capable of either performing, or at least participating in, information processing tasks. For example, Hutchins (1995b) famously speaks about the ways in which the flight-crew and their instruments work together to monitor the speed at which an aircraft is flying; his assertion is that this knowledge does not reside in the head of one specific individual, but is derived from the collection of information that is available in the cockpit. Perhaps, a point to note here is that, ultimately, there needs to be some 'cognizing entity' that is capable of combining the various bits of data into a coherent 'whole' and that this requires a set of mental capabilities that are uniquely human.

Both views raise questions that relate to the manner in which cognition becomes a matter of sharing tasks. In terms of distributed cognition, the work reported in this paper covers both the 'weak' and 'strong' views of distributed cognition. From the 'weak' view, it is argued that the training, knowledge and experience of Crime Scene Examiners allow them to use the environment and the artefacts within it, together with the collection of narratives through the criminal justice process, as resources-for-action in a manner that might be alien to the non-expert. In this way, the Crime Scene Examiner will not only search for specific artefacts but also be able to identify locations which could yield non-visible materials (e.g., places to check for fingerprints, DNA and other evidence). The use of eye-tracking and verbal protocol from crime scene examination show how the approach to searching a scene differs with experience. From the 'strong' view, the reporting and interpretation of evidence from a crime scene through the criminal justice process implies a collective activity (which might not be coordinated by a central agency) that accumulates information to a point at which its interpretation can be tested in Court. While neither approach should be taken to imply that mental states are distributed across individuals, both imply that the action of one individual will form the basis for actions of the next. In this manner, the criminal justice process is able to 'know' the collected evidence, even though it is unlikely that a single individual will have access to all of the information collected during the examination.

Acknowledgements The work reported in this paper was supported by a grant from EPSRC [GR/S85115 MsSAM (Methods to Support Shared Analysis for Mobile Investigators). I am grateful to all of the people who worked with me on this project, and to the Crime Scene Examiners who gave an insight into their work.

References

Artman, H., & Waern, Y. (1999). Distributed cognition in an emergency co-ordination center. *Cognition, Technology and Work, 1,* 237–246.

Artman, H., & Garbis, C. (1998). Situation awareness as distributed cognition. In T. R. G. Green, L. Bannon, C. P. Warren, & J. Buckley (Eds.), *ECCE 9: Proceedings of the Ninth European Conference on Cognitive Ergonomics* (pp. 151–156). Le Chesnay, France: European Association of Cognitive Ergonomics (EACE).

Baber, C., & Butler, M. (2012). Expertise in Crime Scene Examination: Comparing search strategies of expert and novice Crime Scene Examiners in simulated crime scenes. *Human Factors*, *54*, 413–424.

Baber, C., Smith, P., Cross, J., Hunter, J., & McMaster, R. (2006a). Crime scene investigation as distributed cognition. *Pragmatics and Cognition, 14,* 357–385.

Baber, C., Smith, P., Panesar, S., Yang, F., & Cross, J. (2006b). Supporting crime scene investigation. In N. Bryan-Kinns, A. Blandford, P. Curzon, & L. Nigay (Eds.), *People and computers XX—engage* (pp. 103–116). London: Springer.

Bang, M., & Timpka, T. (2003). Cognitive tools in medical teamwork: The spatial arrangement of patient records. *Methods of Information in Medicine, 42,* 331–336.

Busby, J. S. (2001). Practices in design concept selection as distributed cognition. *Cognition, Technology and Work, 3,* 140–149.

Clark, A. (2008). *Supersizing the mind: Embodiment, action and cognitive extension*. Oxford: Oxford University Press.

Doyle, A. C. (1989). *The original illustrated Strand Sherlock Holmes*. Ware: Wordsworth Editions.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine*. New York: The Free Press.

Dror, I. E., Péron, A., Hind, S.-L., & Charlton, D. (2005). When emotions get the better of us: The effect of contextual top-down processing on matching fingerprints. *Applied Cognitive Psychology, 19,* 799–809.

Dror, I., & Harnard, S. (2008). *Cognition distributed: How cognitive technology extends our minds*. Amsterdam: John Benjamin.

Flin, R., Pender, Z., Wujec, L., Grant, V., & Stewart, E. (2007). Police officers' assessment of operational situations. *Policing: An International Journal of Police Strategies and Management, 30,* 310–323.

Flor, N. V., & Hutchins, E. L. (1991). Analyzing distributed cognition in software teams: A case study of team programming during perfective software maintenance. In J. Koenemann-Belliveau, T. G. Moher, & S. P. Robertson (Eds.), *Empirical studies of programmers: fourth workshop* (pp. 36–64). Norwood, New Jersey: Ablex Publishing Corporation.

Furness, D., & Blandford, A. (2006). Understanding emergency medical dispatch in terms of distributed cognition: A case study. *Ergonomics, 49,* 1174–1203.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and Knowing*. Hillsdale, NJ: Erlbaum.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Hillsdale, NJ: Erlbaum.

Hollan, J., Hutchins, E., & Kirsch, D. (2002). Distributed cognition: Toward a new foundation for human-computer interaction. In J. Carroll (Ed.), *Human-computer interaction in the new millennium* (pp. 75–94). New York: Addison-Wesley.

Horswell, J. (2004). *The practice of crime scene investigation*. London: CRC Press.

Hutchins, E., and Klausen, T. (1998). Distributed cognition in an airline cockpit. In Y. Engeström & D. Middleton (Eds.), *Cognition and communication at work* (pp. 15–34). Cambridge: Cambridge University Press.

Hutchins, E. (1995a). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (1995b). How a cockpit remembers its speed. *Cognitive Science, 19,* 265–288.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. In *Human Factors and Ergonomics Society 30th Annual Meeting* (pp. 576–580).

Nemeth, C. (2003). How cognitive artefacts support distributed cognition in acute care. In *Proceedings of The 47th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 381–385). Santa Monica, CA: Human Factors and Ergonomics Society.

Ormerod, T. C., Barrett, E. C., & Taylor, P. J. (2008). Investigating sensemaking in criminal contexts. In J. M. Schraagen, L. G. Militello, T. Ormerod, & R. Lipshitz (Eds.), *Naturalistic decision making and macro-cognition* (pp. 81–102). Avebury: Ashgate.

Perry, M. (2003). Distributed cognition. In J. Carroll (Ed.), *HCI models, theories and frameworks: Towards and interdisciplinary science* (pp. 193–223). San Francisco, CA: Morgan Kauffman.

Putnam, H. (1975). The meaning of meaning. In *Language, mind and knowledge: Philosophical papers* (vol. 2). Cambridge: Cambridge University Press.

Rogers, Y. (1997). *A brief introduction to distributed cognition*. http://mcs.open.ac.uk/yr258/papers/dcog/dcog-brief-intro.pdf.

Schraagen, J. M., & Leijenhorst, H. (2001). Searching for evidence: Knowledge and search strategies used by forensic scientists. In E. Salas & G. Klein (Eds.), *Linking expertise and naturalistic decision making* (pp. 263–274). Mahwah, NJ: LEA.

Seagull, F. J., Plasters, C., Xiao, Y., & Mackenzie, C. F. (2003). Collaborative management of complex coordination systems: Operating room schedule coordination. In *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1521–1525). Santa Monica, CA: Human Factors and Ergonomics Society.

Smith, P. A., Baber, C., Hunter, J., & Butler, M. (2008). Measuring team skills in crime scene examination: Exploring ad hoc teams. *Ergonomics, 51,* 1463–1488.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54,* 277–295.

Chapter 4
Thinking with External Representations

David Kirsh

Abstract Why do people create extra representations to help them make sense of situations, diagrams, illustrations, instructions and problems? The obvious explanation—external representations save internal memory and computation—is only part of the story. I discuss seven ways external representations enhance cognitive power: they change the cost structure of the inferential landscape; they provide a structure that can serve as a shareable object of thought; they create persistent referents; they facilitate re-representation; they are often a more natural representation of structure than mental representations; they facilitate the computation of more explicit encoding of information; they enable the construction of arbitrarily complex structure; and they lower the cost of controlling thought—they help coordinate thought. Jointly, these functions allow people to think more powerfully with external representations than without. They allow us to think the previously unthinkable.

This essay is an inquiry into why thinking and sense making, so often, is interactive. By 'interactive' I mean a back and forth process: a person alters the outside world, the changed world alters the person, and the dynamic continues. Reading a text silently is not an interactive process, for my purposes here, though it is extremely active. Reading and underlining the text, or reading and summarizing it, even reading and moving one's lips, are.

The puzzle that interaction raises about sense making and thinking can be posed like this. In a closed world, consisting of a person and an external representation—a diagram, illustration, spoken instruction, or written problem statement—why do people do more than just think in their heads? If we assume there is no one to ask, no tool to generate novel results, no clock to provide chronometric input, no process to run and observe, then there is nothing external, no oracle or tool, that a person can consult or manipulate, that yields new information. The environment contains

---

D. Kirsh (✉)
Department of Cognitive Science, University of California, San Diego, La Jolla CA 92093-0515, USA
e-mail: kirsh@ucsd.edu

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_4

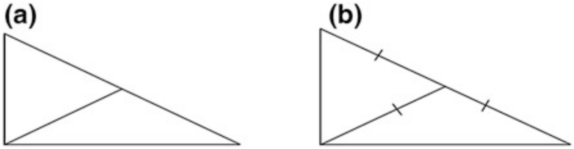


Fig. 4.1 By drawing an example of a right angle triangle and median, it is easier to understand the claim 'in a right-angled triangle the median of the hypotenuse is equal in length to half the hypotenuse'. The illustration does not carry the generality of the linguistic claim, but it is easier to convince ourselves of its truth. In **b** the equalities are explicitly marked and the claim is even easier to read, and helps hint at and helps hint at problem solving approaches

nothing that could not be inferred through reflection, at least in principle. So why bother to make marks, gesture, point, mutter, manipulate inert representation, write notes, annotate, rearrange things, and so on? Why not just sit still and 'think'?

Figure 4.1a illustrates a simple case where interaction is likely. A subject is given the sentence, S_1:

> A basic property of right-angled triangles is that the length of a median extending from the right angle to the hypotenuse is itself one half the length of the hypotenuse.

What do people do to understand S_1? After re-reading it a few times, if they have a good imagination and some knowledge of geometry, they just think. They make sense of it without physically interacting with anything external. Most of us, though, reach for pencil and paper to sketch a simple diagram to help, such as Fig. 4.1a or b. Why? If the sentence were "The soup is boiling over" or "A square measuring 4 in. by 4 in. is larger than one measuring 3 in. by 3 in." virtually no on would bother. Comprehension would be automatic.

Anyone who believes in situated, distributed, or extended cognition will have a ready explanation. Cognitive processes flow to wherever it is cheaper to perform them. The human 'cognitive operating system' extends to states, structures, and processes outside the mind and body. If it is easier to understand a sentence by creating a diagram to help interpret it, then one does that instead of thinking internally alone. The analogy is with a computer system that has memory systems and scratch pads in different media and locations. The decision whether to work out a computation in one or more scratch pads is determined by the type of operators available in each, the cost of operating in each pad, and the availability of space to work in. Processes should migrate to wherever they are best, or most easily, performed.

Figure 4.2 is suggestive of this view of extended or distributed cognition. Because people are embedded in their environments, they are densely coupled to the outside. Cognitive processes drift to wherever they are more cost effective. It's all about the cost structure of computation in each of the interconnected sub-systems. Evidently, when pen and paper is handy, and when the sentence is complex enough, it pays to make a good illustration; it reduces the overall cognitive cost of sense making.

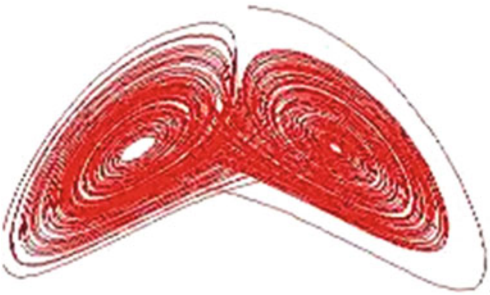


Fig. 4.2 This image of a coupled system represents the state space trajectory over time of certain cognitive processes. Processes readily move from one side to the other, wherever the cost of an operation is lower

Although I believe this is, essentially, a correct account, it is only one of the reasons people interact with external representations. The others have to do with ways changing the terrain of cognition can do more than change cost structure. Chiefly, these involve access to new operators—you can do something outside that you cannot inside; or, you can encode structures of greater complexity than you can inside, external mechanisms allow us to bootstrap to new ideas and new ways of manipulating ideas; or, thirdly, you can run a process with greater precision, faster, and longer outside than inside—you can harness the world to simulate processes that you cannot simulate internally, or cannot simulate as well. In short, these other ways concern changing the domain and range of cognition. This is a striking claim I will justify toward the end.

There is a further reason people interact with external representations: to prepare themselves to coordinate internal and external states, structures, and processes. This feature of interaction is fundamental to our understanding of external representations but rarely studied. See Kirsh (2009a, c). For example, before subjects use a map to wayfind, they typically orient or 'register' the map with their surroundings; they put it into a usable correspondence with the world. Many people also gesture, point, talk aloud, and so on. In principle, none of these actions are necessary to establish a correspondence between elements in the map and the things those elements refer to. Eye movements, mental projection, and other non-interactive techniques may suffice for map-based navigation. But external interactions are commonplace, and a major aspect of understanding representations.

I have found these 'extra' actions also pervasive when people try to understand and follow instructions. In pilot studies, we found that subjects engage in 'interpreting' actions when they follow origami instructions. They register the origami paper with the instruction sheet; they point to elements on the instruction sheet and then focus attention on the counterpart property of the paper; they mutter, they gesture, they move the paper about. This activity is part of processing the meaning of the instructions.

To a lesser degree, the same thing often happens when non-expert cooks follow recipes. They keep place with their finger; they arrange the ingredients to encode their order of use (Kirsh 1995); they read the recipe aloud, ask themselves questions about ingredients, or mutter reminders. We observe similar behavior when people

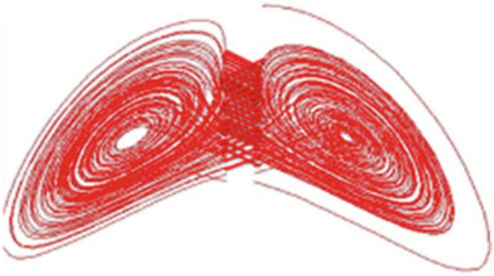


Fig. 4.3 This illustration suggests that there are three cost structures: the cost of inner operations on states, structures, processes, the cost of outer operations on states, structures, processes, and the cost of coordinating inner and outer processes, which includes the cost of anchoring projections, and the cost of controlling what to do, when, and where to do it

assemble furniture. Far from just thinking and then executing instructions, people perform all sorts of apparently ‘superfluous’ actions that facilitate comprehension. They point, mumble, move the instruction manual around, encode order of assembly in the arrangement of pieces. These actions are not incidental, they are often vitally important to sense making and effective action.

One function of these extra actions is to help people anchor their mental processes on external features or processes. Another is to help them tease out consequences, to deepen their semantic and pragmatic processing of the instructions. In both cases, people need to establish a coordination between what goes on inside their heads and what goes on outside. They construct a correspondence, a coordination relation, a synchronization. Because these coordination processes are not cost-free, Fig. 4.2 is overly simple. We have to add to the figure a representation of the coupling process itself, the special actions performed to establish a cognitive link. Figure 4.3 illustrates this added cost-laden process: anchoring (see Kirsh 2009b, for an initial discussion of this third cost space).

As important as the anchoring—or grounding process—is I restrict my focus, in the remainder of this work, to ways we interact with representations to alter the cognitive terrain rather than the interactions we perform to prepare ourselves to engage the external part of that terrain through anchoring.

Materiality and Its Consequences

The argument others and I have long advanced is that people interact and create external structure when thinking because:

> Through interaction it is easier to process more *efficiently* and more *effectively* than by working inside the head alone (Clark 2008; Kirsh 1995, 1996, 2009a, b, c; Kirsh and Maglio 1994).

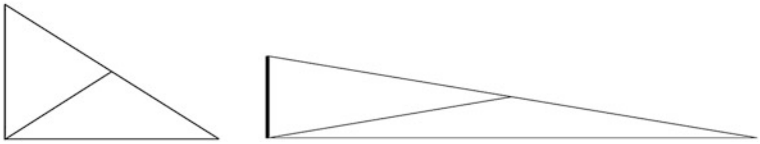


Fig. 4.4 Choices must be made when drawing a triangle. Should the triangle be long and short? Isosceles? Will any of these choices affect the truth of the sentence? By having to resolve these questions, subjects are helped in the problem solving process

Efficiency usually translates as speed-accuracy. Interactive cognition regularly leads to fewer errors or to greater speed. Effectiveness usually translates as coping with harder problems. Interactive cognition regularly helps subjects to compute more deeply, more precisely, and often more broadly.

The idea is that by operating with external *material*, with pen, paper, ruler, and then working to meet one's goals and sub-goals using that external material—*draw* a triangle, *mark* the half point of the hypotenuse—subjects benefit from physical constraint and visual hints that help cognition. This plays out in a few ways. For instance, the constructive process helps drive interpretation. Because action is primarily serial, it is incremental; a structure emerges step-by-step and a subject must resolve specific problems. What size should the base and height be? Does it matter? Does the median bisect the right angle? Working with tools and the external structure, moreover, grounds interpretation in an ever more constrained case study. After choosing the size of the right angle triangle, the requirement to split the hypotenuse in half is very concrete. It is now 'split this hypotenuse'. This incremental, interactive process, filled with prompts, hints, visible possibilities and impossibilities, provides more constraint than mentally computing a conceptual whole solely from the semantics of linguistic parts. The linguistic formulation is more general, but it is also less constrained. See Figs. 4.1 and 4.4.

A second way materiality figures in cognition is by explicitly involving visual and motor cortex. When a structure is viewable and drawable, its properties prime a constellation of associations. Just by grappling with external material—using rulers, making lines intersect—and then looking at the results, a set of properties and possibilities of forms are encountered and primed. For instance, if two lines intersect then they define a set of angles. It is natural for visual attention to focus on estimating angles. Are they equivalent? If the triangle has a right angle, then automatically a network of spatial concepts related to right triangles are activated, particularly associations derived from previous work with diagrams of right triangles. These visual and physical associations may be different and more extensive than associations derived from verbal accounts. This is apparent whenever a tool is in hand. Rulers prime measuring actions and thoughts; protractors encourage thoughts of angles and degrees.

The benefits of interacting with an external representation are especially clear for complex structures. As the complexity of a linguistic specification of a visual structure increases, it becomes more rewarding to make sense of the sentence by constructing a physical drawing and looking at it, than by constructing that geometric form in one's mind's eye and making sense of the sentence internally. Most people find it easier to think in terms of physical lines than in terms of the mental counterparts of lines, particularly the more lines there are, or the more complex the structure. Even though some people can do things in their heads that others cannot, there is always a point where internalist cognitive powers are overwhelmed and physical realization is advantageous (see Kirsh 2009b). Thus, although from a purely logical point of view, a closed system of world and person contains no additional information after that person has drawn an interpretation than before, there nonetheless are important changes wrought by interaction that can positively alter the cognitive terrain. Specifically, these interactive changes concern:

- What's *active inside* the person's head—what's being attended to, what's stored in visual or motor memory, and what's primed—an external structure encourages a visual scanpath that activates expectations, drawing the structure displays angles, lengths, and will cause distant cognitive associations in motor and visual cortex;
- What's *persistent outside*, and in the visual or tangible field—an external structure holds a structure constant until it is added to; the structure does not decay the way mental structures and processes do, and it supports repeated perceptual inquisition;
- How information is encoded, both inside and outside in virtue of interaction. Because there is an external structure present, subjects can try out different *internal and external representational forms*, the two forms can play off each other in an interactive manner, leading to new insights.

The upshot is that, often, humans are able to improve their thinking and comprehension by creating and using external representations and structures. By working outside, they change what is inside and interactively they can reach new thoughts. This may be stunningly obvious, yet it is sufficiently foundational and far-reaching to deserve analytic and empirical exploration.

Let me press this idea further by turning now to seven distinct benefits that externalization of structure confers.

Shareable and Identifiable Objects of Thought

When someone externalizes a structure, they are communicating with themselves, as well as making it possible for others to share with them a common focus. An externalized structure can be shared as an object of thought. This reification of internal object—this externalization—has benefits for both parties.

Here is an example. In Fig. 4.5, an explicit geometric form has been added to the body position of a dancer. Using a video to demonstrate torsion, Bill Forsythe, a noted choreographer, had his colleagues visually annotate key body features on the

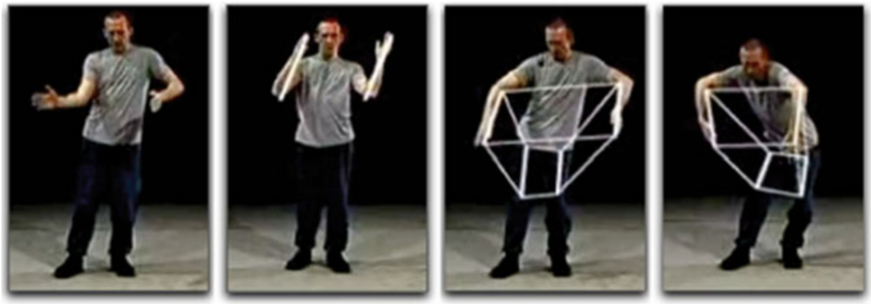


Fig. 4.5 Bill Forsythe, a noted contemporary choreographer, has begun documenting certain concepts and principles of choreography in film. Here he explains torsion. The annotation makes it easy for the audience to refer to otherwise invisible structures

video as he spoke. He first identified points on his body, orally, then, as he turned his discussion to line segments, such as the line between elbow and hand, these were superimposed on the video, and finally he talked of joining the segments into a three dimensional trapezoid, and his viewers saw a representation of the three dimensional form appear on screen. It was then easy for viewers to see the effect of movement on deformations of the trapezoid. Forsythe relied on his listeners seeing the visible annotation, the trapezoidal structure, as he explained the ideas of torsion, sheer, and body axes (Forsythe 2008).

One virtue of this particular annotation is that by having verbally defined the structure to be manipulated, and then visibly locating it on his body, the choreographer and anyone looking at the video, knows that if they refer to any visible part of the trapezoid their reference will be understood. They can ask pointed questions about how the shape figures in what the speaker is saying, or even how some specific feature—the apex or base—figures in an abstract idea. For instance, once there are external lines and planes anyone can ask the speaker, or themselves, which body positions keep the volume of the shape constant, or which movements ensure the top plane remains parallel to the bottom plane. Choreographers find such questions helpful when thinking about body dynamics and when they want to communicate ideas of shearing and torsion to their dancers. But they are hard to understand if the group does not share a visual or projected image of a transforming shape.

Physically reifying a shape through annotation adds something more than just providing a shared reference; it provides a persistent element that can be measured and reliably identified and re-identified. Measurement is something one does after a line or structure has been identified. This need not always require an external presence. Some people are able to grasp the structure of a superimposed trapezoid purely by mentally projecting an invisible structure onto the body. They listen to the speaker; watch his gestures, and project. But even for these strong visualizers, annotating still helps because once something is externalized it has affordances that are not literally present when projected alone.

For instance, when the lines of a shape are externalized we can ask about the length of the segments and their angles of intersection. We know how to measure these elements using ruler and protractor. Lines afford measuring. Granted, it is still possible, though not easy, to measure the length of mentally projected lines if the subject is able to appropriately anchor the projected lines to visible points. A choreographer, for instance, can refer to the length of someone's forearm through language or gesturally mark a structure without having to annotate a video. But can he or she refer reliably to the length of lines *connecting* the top and bottom planes of a complex structure without having those planes visibly present? Those lines have to be anchored on the body. If a structure is as complex as a truncated pyramid, which has eight anchor points, it must be constructed in an orderly manner, much as Forsythe did in his annotated video, else there is too much to keep track of mentally. This does not decisively show that such structures cannot be identified and marked out by gesture and posture without visible annotation. But the complexity of mental imagery and mental projection goes way up as the number of anchors increases; or when the target body moves the anchor points; or, worst of all, if invisible anchor points are required, as would be the case if the conceptualized pyramid were to extend right up to its apex. The peak itself would be floating in air, unconnected to anything material, anchorless. Imagine trying to use that invisible anchor as an anchor for something else. By contrast, once the form is made manifest in visible lines, all such elements can be explicitly referred to, even visibly labeled; they can be located, measured, intentionally distorted if so desired, and the nature of their deformation over time can be considered. They become shared objects of thought.

This is worth elaborating. To say that something is, or could be, an object of thought implies the thinker can mentally refer to it—in some sense the thinker can grasp the referent. A shared object of thought means that different thinkers share mechanisms of reference and for agreeing on attributes of the referent. For instance, Quine (1960), following Strawson (1959), argued that objects must have identity conditions, as in his motto "No entity without identity". Entities have to be identifiable, re-identifiable, and individuatable from close cousins. Would the structures and annotations in Fig. 4.5 meet those criteria if imagined or projected mentally? It depends on how well they are anchored to physical attributes. Certainly there are some people—choreographers, dancers, and people with wonderful imaging abilities—who can hold clear ideas of projected structure, and use them to think with. As long as there is enough stability in the 'material anchors' (Hutchins 2005) and enough expertise among the subjects to ensure a robust projection, the lines and shapes these experts project onto the visible environment meet most criteria of 'entification', though, of course this is purely an empirical claim. But most of us find that it is easier to think about a structure that has been reified by adding visible or tangible elements to the environment. The structure is more vivid, more robust, clearer—a better object of thought. Almost everyone needs to see the lines and shapes to see subtle geometric relations between them. So, we create external structure. It is by this act of materializing our initial projections, by forming traces of those projections through action, or material change, that we create something that can serve as a stepping-stone for our next thoughts.

This interactive process of *projecting structure then materializing it*, is, in my opinion, one of the most fundamental processes of thought. When we interact with our environment for epistemic reasons, we often interact to create scaffolds for thought, thought supports we can lean on. But we also create external elements that can actually serve as vehicles for thoughts. We use them as things to think with.

All too often, the extraordinary value of externalization and interaction is reduced to a boring claim about external memory. “Isn’t all this just about offloading memory?” This hugely downplays what is going on. Everyone knows it is useful to get things out of the head and put where they can be accessed easily any time. It is well known that by writing down inferences, or interim thoughts, we are relieved of the need to keep everything we generate active in memory. As long as the same information can be observed and retrieved outside, then externalizing thought and structure does indeed save us from tying up working memory and active referential memory.

But memory and perception are not the same thing. Treating information to be the same whether it is outside or inside ignores the medium specific nature of encoding. The current view in psychology is that when we visually perceive an external structure, the information that enters is stored first in visuo-spatial store (Baddeley 2000; Logie 1995), before being processed for use in later mental processes. Since the form a structure is encoded in profoundly affects how easily it can be used in a process, it is an open question how much internal processing is necessary to convert an external structure into an internal structure that is usable. Accordingly, it cannot be assumed, without argument, that the costs are always lower in perceptually retrieving information than ‘internally’ retrieving information, even if that information is complex and voluminous and something we would normally assume is more efficiently stored externally. The strength of this concern is obvious if the information element to be perceived is buried in visual clutter. Much will depend on visual complexity, the form information is encoded in, how easy it is to perceive the structure when it is wanted, and so on. Even when an object of thought is present in a clear and distinct way—as Forsythe’s graphical annotations are—it still must be perceived, then gestalted, and conceptualized. Do we really know the relative cost of grasping an externally represented content versus an internally represented one?

The implication is that using the world as external storage may be less important as a pure source of cognitive power than using the world for external computation. Things in the world behave differently than things in the mind. For example, external representations are extended in space, not just in time. They can be operated on in different ways; they can be manually duplicated, and rearranged. They can be shared with other people. Tools can be applied to them. These differences between internal and external representations are incredibly significant. They are what makes interactivity so interesting.

I turn now to another of these differences: the possibility of manually reordering physical tokens of statements. Because of rearrangement, it is possible to discover aspects of meaning and significance—implications—that are hard to detect from an original statement when viewed in isolation. By reordering and rearranging what is close to what, we change a token's neighborhood, we change the space of what is cognitively near.

Rearrangement

The power of physical rearrangement, at least for vehicles of propositions, such as sentences, logical formulae, pictorial narratives, is that it lets us visually compare statements written later with those written earlier; it let's us manipulate what is beside what, making it easier to perceive semantically relevant relations. For instance, we can take lemmas that are non-local in inference space—inferences that are logically downstream from the givens and usually discovered later, hence written further down the page—and rewrite them so they are now close to earlier statements. Statements that are distant in logical space can be brought beside each other in physical space. If we then introduce abbreviations or definitions to stand in for clusters of statements, we can increase still further the range of statements we can visually relate. This process of inferring, duplicating, substituting, reformulating, rearranging and redefining, is the mechanism behind proofs, levels of abstraction, the lisp programming language, and indeed symbolic computation more generally.

The power of *rearrangement* is shown in Fig. 4.6. The problem is to determine whether the six pieces on the left are sufficient to build the form on the right. What do you need to do to convince yourself? Since the problem is well posed and self contained, the question again, is 'why not just work things out in your mind?' In Fig. 4.6, you have no choice: because the pieces are not movable, no doubt, you will confine your thinking to looking and imagining the consequences of moving and rotating them. But, if the problem were posed more tangibly, as a jigsaw puzzle with movable tiles, wouldn't it be easier to try to *construct* an answer in the world than to think through an answer internally?

Reorganizing pieces in physical space makes it possible to examine relations that before were distant or visually complex (e.g., rotations and joins). By re-assembling the pieces, the decision is simply a matter of determining whether the pieces fit

Fig. 4.6 Can the jigsaw images on the left be perfectly assembled into the picture on the right? If you could rearrange the pieces, the answer would be trivial. The answer is no. Can you see why? Why, in general, is it easier to solve jigsaw puzzles tangibly?

perfectly together. That is a question resolvable by physically fitting and visually checking. Interaction has thus converted the world from a place where internal computation was required to solve the problem to one where the relevant property can be perceived or physically discovered. Action and vision have been substituted for imagery, projection, and memory. Physical movement has replaced mental computation. Instead of imagining transformations, we execute them externally.

It is tempting to interpret the benefits of rearrangement entirely in cost structure terms: processes migrate to the world because they are cheaper or more reliable there. Evidently, physical manipulation, at times, is cognitively more efficient and effective than mental manipulation. So, on those occasions, it is rational to compute externally.

And sometimes that is all there is to it. For example, in Tetris, subjects can choose between rotating a tetrazoid in their heads and rotating it in the world (Kirsh and Maglio 1995). Since physical rotation is, in fact, a bit faster than mental rotation, the cost incurred by occasionally over-rotating a piece in the world is more than made up for by the benefits that come from the faster and less error prone decision making based on vision.

Yet, it is not always so. In solving jigsaw puzzles, more is at stake than cost alone. As the descriptive complexity of the board state increases, there comes a point where it is hard, if not impossible, for someone to hold the complete structure in mind. The very act of trying out a move mentally causes the internally maintained structure to degrade. Imagine trying to assemble twenty separate pieces in your mind, and then checking to see if the twenty-first will fit anywhere in the mentally sustained assembly. The twenty-first piece may be the last straw, total overload, causing the whole mental structure to lose integrity.

The analogy is with swap space in a computer. Once a threshold of complexity is reached, a computer begins to degrade in its performance. In fact, if its flailing is serious enough, it reaches a standstill, where it takes so much of its working memory to hold the board state, that the simple act of changing that state exhausts memory. The system lacks the resources to keep track of what it has tried already and what remains to be tried. It has to place in long term memory the part of the board state it is not currently checking, so that it can process the steps in its program telling it what to do next. Then, to do the next thing, it has to bring back part of the board state in long-term memory, and swap out the control state. The result is that the system may cycle endlessly in the same subset of states, never canvassing the part of the state space where the solution is to be found. Zero progress.

It is not quite like that in the world. Because of physical persistence, the board remains the same before and after a subject thinks about moves. Unlike the mental realm, the stability of a physical state is not significantly affected by its complexity. A twenty-piece assemblage is just as stable as a ten-piece assemblage.

There are limits in the physical world too. Once a board arrangement has been changed physically, the previous state is lost, unless a further trace, an annotation was created, or a digital image taken. So searching for a solution in the world, as opposed to in the head, is not always better. But with enough externalization of state—enough external record keeping—there are jigsaw puzzles that can be solved physically that would be impossible to solve in the head, through mental simulation alone. We can push the complexity envelope arbitrarily far. This cannot be done in the head alone.

I will return to this topic of in principle differences between mental and physical simulation at the end of the essay.

Physical Persistence and Independence

Both rearrangement and having stable objects to think with both rely on physical things being persistent. The next key difference between internal and external representations, then, concerns the difference in their stability and persistence over time. Rearrangement of jigsaw pieces is possible because the different pieces to be arranged are simultaneously present. If six pieces were present before rearrangement, there are six after. Pieces can be moved nearer to each other without destroying their integrity. Even though things are not quite as simple with thinking with physical tokens of sentences, we still can be confident that we have the same thought before and after moving a sentence. Because it is easy to detect differences in a sentence token simply by comparing the original with its copy—that is, before and after copying a sentence inscription—we depend on physical persistence to ensure we do not change the object of thought just by copying or moving tokens. Other things equal, the sentence ‘this sentence has five words’ means the same whether printed on the right or the left of the page, and whether printed yesterday or today.

The case is rather different for mental representations. How can a subject be sure that the mental image in mind at time t_1 is the same as the one at t_0? And how can a subject know whether the addition of another mental image, or a simple rotation of a mental image, has not changed the original image? The only reliable test is whether the image is caused by the same external structure on both occasions. If that structure is not present, there is no objective touchstone to decide sameness. There is just subjective belief. For Wittgenstein (1953) this was a source of skepticism concerning the possibility of knowing one’s mental state without outside referents to ground it. No inner state or inner process without outer criterion. Hence, without external support, there might be no way of knowing whether one has the same thought on two occasions.

The brute fact of physical persistence, then, changes the reliability, the shareability, and the temporal dynamics of thinking. It is easier to have the same thought tomorrow, if the vehicles encoding the thought, or the cues stimulating the thought, are the same as today’s. That’s why writing helps. When the vehicle is external we can also count on other people ratifying that it remains the same over time. So, we can be confident that if we think we are reading the same sentence on two occasions, there is a fact of the matter. Similarly, we can be confident that if we interact with an external representation and we think we have left it unchanged, our judgments are more reliable than those concerning our beliefs about internal representations. Moreover, in the outside world, there is widespread empirical agreement on the effect of interaction—we know there is a broad class of transformations that leave structures invariant, for example: rotation, translation, lighting change, and so forth. There is no comparable principle for internal representations. We have no way of knowing the constancy of our inner life. This means that we have a better idea of the effect of interacting with external representations than with internal ones.

Physical persistence also differs from mental persistence, and transient mental presence, in increasing the *range of actions* a subject can perform on the underlying thing encoding the representation—the vehicle. In Fig. 4.5, for example, the truncated '3D' trapezoid is displayed as a line drawing on the choreographer's body. It is shown in stop action. Measurements can be made because the visible structure—the trapezoid—can be frozen for as long as it takes to perform the measurements. Tools can be deployed. The materiality of external representations provides affordances internal representations lack.

Architects, designers, and engineers exploit the benefits of persistence and material affordance when they build models. Models have a special role in thinking, and can for our purposes be seen as two, three, or even four-dimensional external representations: paper sketches—2D; cardboard models, cartoons, and fly throughs—3D models in space or time; and dynamically changing three dimensional spatial structures—4D models. To see the extra power these sort of external representations offer let us look at the scale models that architects build.

Scale models are tangible representations of an intended design. They serve several functions:

1. They can serve as a shared object of thought because they are logically and physically *independent* from their author. They can be manipulated, probed, and observed independently of their author's prior notion about how to interact with the model. This is vital for talking with clients, displaying behavior, functionality and detecting unanticipated side effects. It makes them public and intersubjective.
2. Models enforce consistency. The assumption behind model theory in mathematics is that if a physical structure can be found or constructed, the axioms that it instantiates must be consistent (Nagel and Newman 1958). Unlike a description of the world, or a mental representation, any actual physical model must be self-consistent. It cannot refer to properties that are not simultaneously realizable, because if it is a valid model it counts as an existence proof of consistency. In a many part system, part A cannot be inconsistent with part B if they both can simultaneously be present in the same superstructure. Similarly, the movement of part A cannot be inconsistent with the movement of part B if the two can be run simultaneously. Build it, run it, and thereby prove it is possible. Inconsistency is physically unrealizable. There are few more powerful ideas than this.
3. Models reveal unanticipated consequences. To say that an external model is independent of its creator is to emphasize that other people can approach the

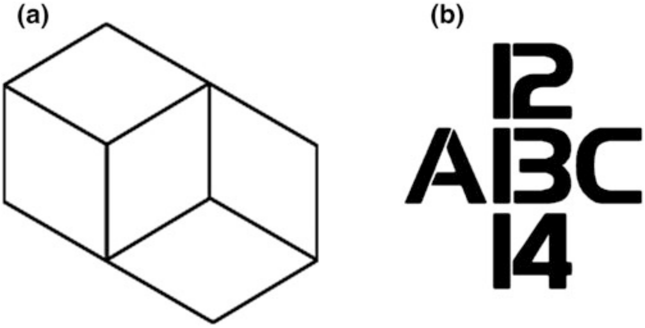


Fig. 4.7 Here are some ambiguous objects. In **a**, a variant of the Necker cube is shown where some corners that start looking convex (outward pointing) change to concave. **a** is ambiguous in several ways. Can you see at least four interpretations? In **b**, the middle element will seem to be a B or 13 depending on whether you read vertically or horizontally. How an object visually appears depends on how an agent looks at it, and this can be affected by how the structure is framed, how it is contextualized, how the agent feels, or what an agent is primed to see

model in ways unconstrained by its creator's intention. Once a structure is in the public domain it has a life if its own. This is well appreciated in the case of ambiguous objects. Look at Fig. 4.7a. Its author may have intended it to be a convex cube with two concave sides extending to the bottom right. But a viewer may initially see those concave sides as a convex cube with a corner pointing outward. Look at the image longer and other interpretations should appear. Studies on mental imagery have shown that subjects who have not yet detected an ambiguity by the time they create a mental image are not likely to realize the ambiguity inherent in their image (Chambers and Reisberg 1985). It is as if they are sustaining their image under an interpretation, a prior conception. And so, they are closed to new interpretations. When externalized and in the visual field, however, the very processes of vision—the way the eye moves and checks for consistency—typically drives them to see the ambiguity.[1] When a structure is probed deeply enough, relations or interactions between parts, that were never anticipated may be easy to discover. Thus, an author may be able to discover interpretations he or she never considered. Whether the thing externalized is a representation of a thought, image, or mental animation, its persistence and independence means that it may be reconsidered in a new light, and interacted with in a new manner.

The power of modeling is a topic of its own. Another special property is one that is made explicit in mathematical simulations that can be run back and forth under a user's control. Such simulations provide persistence and author independence because they can be run forward, slowed down, stopped, or compared snapshot by

---

[1] Although this argument concerns visual processing, it applies equally well to the physical interactions we can perform on the physical object.

snapshot. All normal sized models support our physical interaction. We can move them in ways that exposes otherwise hard to see perspectives and relations. When our interaction is controlled precisely, or interpreted as movement along a timeline, we can juxtapose snapshots in time for comparisons that would simply be impossible otherwise. Without the stability of reproducibility and persistence, some of the ideas we form about the temporal dynamics of a structure would be virtually unthinkable.

Reformulation and Explicitness

A fourth source of the power of interaction relies on our ability to externally restate ideas. Sometimes it is easier to perform restatement externally than in our heads.

Representations encode information. Some forms encode their information more *explicitly* than others (Kirsh 1992). For example, the numerals '$\sqrt{2209}$' and '47' both refer to the number 47 but the numeral '47' is a more explicit encoding of 47. Much external activity can be interpreted as converting expressions into more explicit formulations, which in turn makes it easier to 'grasp' the content they encode. This is a major method for solving problems. For instance, the problem $x = \sqrt{4}28,561 + \sqrt{2209}$ is trivial to solve once the appropriate values for $\sqrt{4}28,561$ and $\sqrt{2209}$ have been substituted, as in x = 13 + 47.[2]

Much cognition can be understood as a type of external epistemic activity. If this seems to grant the theory of extended mind (Clark 2008) too much support add the word 'managing' as in 'much cognition involves managing external epistemic activity'. We reformulate and substitute representations in an effort to make content more explicit. We work on problems until their answer becomes apparent.

The activity of reformulating external representations until they encode content more transparently, more explicitly, is one of the more useful things we do outside our heads. But why bother? Why not do all the reformulation internally? A reason to compute outside the head is that outside there are public algorithms and special artifacts available for encoding and computing. The cost structure of computation is very different outside than inside. Try calculating $\sqrt{2209}$ in your head without relying on a calculator or an algorithm. Even savants who do this 'by just thinking' find there is a limit on size. Eventually, whoever you are, problems are too big or too hard to do in the head. External algorithms provide a mechanism for manipulating external symbols that makes the process manageable. Indeed were we to display the computational cost profiles (measured in terms of speed accuracy) for performing a calculation such as adding numbers in the head versus using

---

[2] Reformulation is not limited to formal problem solving. The statement "Police police police police police" is easier to understand when restated at "Police who are policed by police, also police other police". Most people would not break out their pens to make sense of that statement, but few of us can make sense of it without saying the sentence out loud several times.

algorithms or tools in the world, it would be clear why most young people can no longer do much arithmetic in their heads. Tools reshape the cost structure of task performance, and people adapt by becoming dependent on those tools.

A second reason we compute outside rather than inside has to do with a different sort of complexity. One of the techniques of reformulation involves substitution and rewriting. For instance, if asked to find the values of x given that x2 + 6x = 7, it is easiest if we substitute (x + 3)2−9 for x2 + 6x. This is a clever trick requiring insight. Someone had to notice that (x + 3)2 = x2 + 6x + 9, which is awfully close to x2 + 6x = 7. By substituting we get (x + 3)2 = 16, which yields x = 1 or −7. Could such substitutions be done in memory? Not likely. Again, there are probably some people who can do them. But again, there always comes a point, where the requisite substitutions are too complex to anticipate the outcome 'just by thinking' in one's head. The new expressions have to be plugged in externally, much like when we swap a new part for an old one in a car engine and then run the engine to see if everything works. Without actually testing things in the physical world it's too hard and error prone to predict downstream effects. Interactions and side effects are always possible. The same holds when the rules governing reformulation are based on rewrite rules. The revisions and interactions soon become too complex to expect anyone to detect or remember them.

Natural Encoding

Persistence, reordering, and reformulation largely explain why externalizing information and representation may increase the efficiency, precision, complexity and depth of cognition. And if these aspects of interaction with external representations do not explain the extra power to be had then simulation does. Still, there is another aspect to consider: how external processes may increase the *breadth* of cognition. To explore this aspect consider again, why we prefer one modality to another for certain types of thinking (Fig. 4.8).

Every representational system or modality has its strengths and weaknesses. An inference or attribute that is obvious in one system may be non-obvious in another. Consider Fig. 4.9—a musical notation. The referent of the notation is a piece of music. Music is sound with a specific pitch or harmony, volume, timber, and temporal dynamics. The 'home' domain of music, therefore, is sound. Visual notation for music is parasitic on the structure of sound. Prima facie, the best representation to make sense of musical structure is music itself; we go to the source to understand its structure[3].

---

[3] To see why music can be both referent and representation (terrain and also map) ask whether there is a difference between hearing sound and hearing sound **as** music. The sound is the terrain; music is the conceptualizing structure that interprets the sound; it maps it.

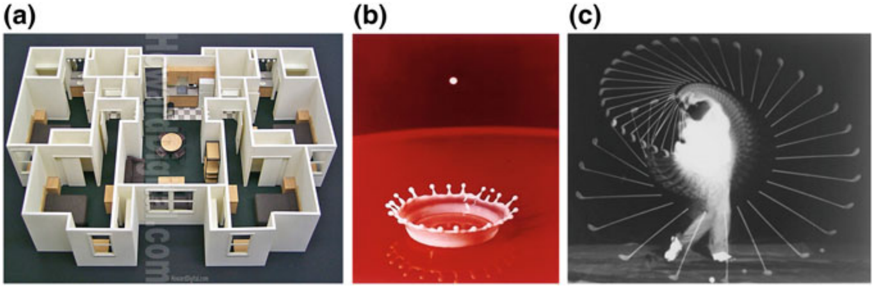


Fig. 4.8 A 3D model permits architects to view a form from arbitrary angles. It allows them to measure, compare, and look for violation of constraints. By approaching the model from odd angles they can see occlusions, and relations that would be extremely hard to see otherwise. In **b** we see a near perfect coronet formed by a drop of milk in a famous photograph by Harold Edgerton. And in **c** we see a famous stop frame image of a golfer swinging and then hitting a golf ball (Densmore Shute Bends the Shaft, 1938, © Dr Harold Edgerton, Silver Gelatin Print)

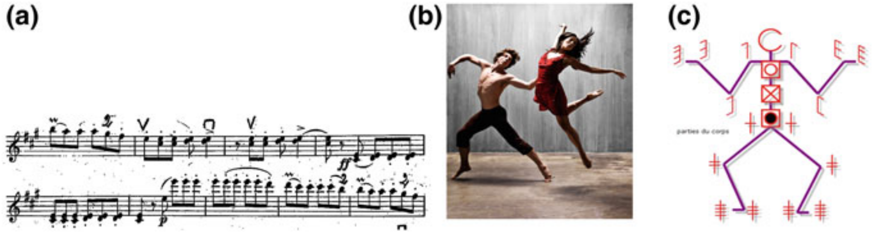


Fig. 4.9 Imagine hearing 12 s of music. Now look at the musical notation shown here. Notation has the value of showing in space a structure that one hears. But there is much more in the sound as heard than is represented in the notation alone. Sound is the natural representation of music. The same is true for dance. Compare Laban notation for dance with the full body structure of dancers. Even if the joint structure is captured in Laban, how well represented are the dynamics of movement, the feel of the dance, and its aesthetic impression?

If there are times when the source medium is required to represent the content of a thought, then a further reason to externalize content and manipulate it outside is that for some problems, the natural representation of the content only exists outside. Arguably, no one—or at best only a few people—can hear music in their head the way it sounds outside. Mental images of sounds have different properties than actual sounds. Even if it is possible for the experience of the mental image of music to be as vivid and detailed as perception of the real thing, few people—other than the musically gifted, the professional musician, or composer (Sacks 2008)—can accurately *control* musical images in their heads. It is far easier to manifest music externally than it is to do so internally. So, for most people, to make sense of music the first thing to do is to play it or listen to it.

This raises a further requirement on the elements of thought. If a representational system is to function as a medium of thought the elements in the system must be sufficiently manipulable to be worked with quickly. Spoken and written words are malleable and fast. Body movements for dance, gesture, and perhaps the pliability of clay are too. Musical instruments, likewise, permit rapid production of sound. These outer media or tools for creating media support fast work. They enable us to work with plastic media. In this respect they enable us to work outside much like the way we work inside, using visual or auditory images for words or ideas, which most of us work with at the speed of thought. If external manipulability matches the internal requirements on speed, then an external medium has the plasticity to be a candidate for thinking in.

Using Multiple Representations

Despite the value of listening to music there are times when notation does reveal more than the music one has listened to—instances where a non-natural representation can be more revealing and intuitive than the original representation. Because a notational representation uses persistent, space consuming representations, early and later structures can be compared, superimposed and transformed using notation specific operators. As with logic and jigsaw puzzles it is useful to have tangible representatives that can be manipulated. In these cases, a subject who moves from one representation to the other may extend cognition. By moving between listening to music, and writing it down in a notation, or listening and then reading the notation, or sometimes vice versa, a composer or listener may be able to explore certain elements of musical structure that are otherwise inaccessible. The more complicated the structure of the music the more this seems to be true. Without interacting with *multiple representations* certain discoveries would simply be out of reach. Visual designers who move between pen and paper, 3D mockups and rapid prototypes are familiar with the same type of process.

Construction and Tools

The final virtue of external interaction I will discuss is, in some ways, the summation of persistence, rearrangement, and reformulation. It may be called the power of construction. In making a construction—whether it be the graphical layovers of the dancer shown in Fig. 4.5, the geometric construction of Fig. 4.1, or building a prototype of a design as in Fig. 4.8a—there is magic in actually making something in the world. As mentioned in the discussion of scale models, by constructing a structure we prove that its parts are mutually consistent. If we can build it, then it must be logically and physically viable. If we can run it, then the actions of those parts are consistent, at least some of the time; and if we can run it under all orderings then it is consistent all of the time. The physical world does not lie.

The constructive process has a special place in human thinking because it is self-certifying. In mathematics, constructive reasoning means proving a mathematical object exists by showing it. For example, if it were claimed that a given set has a largest element, then a constructionist proof would provide a method for finding the largest element, and then apply the method to actually display the element.

Not every form of human reasoning is constructive. Humans reason by analogy, by induction, they offer explanations, and they think while they perform other activities, such as following instructions, interpreting a foreign language, and so on. None of these are constructive methods in the mathematical sense. However, because of the incremental nature of construction the effort to construct a solution may also be a way of exploring a problem. When students look for a constructive proof to a geometric problem, they use the evolving external structure to prompt ideas, bump into constraints, and realize possibilities. When they write down partial translations of a paragraph they rely on explicit fragments to help guide current translation.

The question that begs to be asked is whether thinking with external elements is ever necessary. Can we, in principle, do everything in our heads, or do we need to interact with something outside ourselves in order to probe and conceptualize, and get things right? In mathematics, externalization is necessary, not just for communication, but to display the mathematical object in question. It is like measurement: you cannot provide the value of a physical magnitude without measuring it. You cannot show the reality of a mathematical object (for constructivists) without revealing a proof that parades it. Yet, during the discovery process might not all the thinking be internal, the result of an interaction between elements inside the head? Where is the proof that, at first, all that probing and conceptualizing might not be the outcome of a purely internal activity? Even if the internal activity is simulating what it would be like to write things down outside, or how one would present one's idea to others, all the 'real' thinking lives internally. We needed the outside world to teach us how to think,[4] but once we know how we never need to physically encounter tangible two or three-dimensional structures to epistemically probe the 'world'.

I believe this is wrong: physical interaction with tangible elements is a necessary part of our thinking process because there are occasions when we must harness physical processes to formulate and transition between thoughts. There are cognitive things we can do outside our heads that we simply cannot do inside. On those

---

[4] Vygotsky among others has suggested that we mastered thinking externally, by conforming our behavior to social norms of rational inquiry, and that what we learned to do first on the outside we came to do on the inside. Thus, the reason we can do math in our head is because we can do math in the world. The same applies to thinking internally in auditory images. We think in words internally, using auditory images of sounds, because when we think in public we speak. Thinking internally is simulating what we do externally, though Vygotsky did believe that inner speech of adults would be much compressed and unintelligible to anyone except the thinker (Vygotsky 1986).

occasions, external processes function as special cognitive artifacts[5] that we are incapable of simulating internally.

To defend this hypothesis is harder than it might seem. In practice, few people can multiply two four-digit numbers in their heads. And, if they can, then increase the problem to ten digit numbers. This in practical limitation does not prove the 'in principle' claim, however, that normal human brains lack the capacity to solve certain problems internally that they can solve with external help, with tools, computers or other people. There are chess masters who can play equally well, or nearly as well, blindfolded as open eyed (Chabris and Hearst 2003). There is no evidence that a team of chess players is better than an individual.There are people with savant syndrome who can multiply large numbers in their head, or determine primes or square roots. Other savants with eidetic memories can read books at a rate of 8–10 s per page, memorizing almost everything.[6] Tesla said that when he was designing a device, he would run a simulation of it in his head for a few weeks to see which parts were most subject to wear (Hegarty 2004, p. 281, citing Shepherd). Stephen Hawking is said to have developed analytical abilities that allowed him to manipulate equations in mind equivalent to more than a page of handwritten manipulations. For any reasoning problem of complexity *n*, how do we know there is not some person, somewhere, who can solve it in their head, or could, if trained long enough? To be sure, this says little about the average person. Any given person may reach their computational limit on problems much smaller than *n*. And our technology and culture has evolved to support the majority of people. So, in practice, all people rely on available tools, practices, and techniques for reasoning. Nonetheless, if a single person can cope with *n*, then there is an existence proof that the complexity of external simulation does not itself mean that internal simulation is not possible. It suggests that any problem we cannot solve in our heads that we can solve with external help, has more to do with cost structure than with an in principle biological inability.

One way of making the in principle case is to show that there are *operations* that can be performed on external representations that cannot be performed on internal representations, and that, somehow, these are essential. Are there epistemic activities we can perform outside that we cannot duplicate inside, not because of their complexity, but because there are physical properties and technologies available on the outside that we cannot duplicate mentally—operations we cannot mentally simulate with sufficient realism to deliver dependable answers?

Consider Fig. 4.10. The dots in the two images on the left are related to one another by a rotation of 4°. This is essentially invisible unless the two images are

---

[5] Hutchins (2001).

[6] Entry from Wikipedia on Kim Peek the inspiration for the character in the movie *Rain Man*: "He reads a book in about an hour, and remembers almost everything he has read (…) His reading technique consists of reading the left page with his left eye and the right page with his right eye and in this way can read two pages at time with a rate of about 8–10 s per page. He can recall the content of some 12,000 books from memory. Wikipedia on Kim Peek. http://en.wikipedia.org/wiki/Kim_Peek Nov 2009.

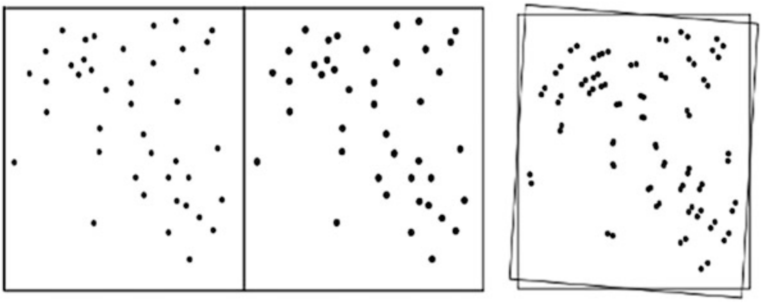


Fig. 4.10 On the left are two collections of random dots. They differ only by the rotation of the plane they are in. On the right, they have been superimposed. Their global relation is now visible. Could this relationship be detected without physically superimposing the patterns? Mental imagery does not support vivid superimposition. And if there are outlier humans who have this odd ability they will necessarily fail as the number of dots or the number of superimpositions increase

superimposed, as in the image on the right. Superimposition is a physical relation that can be repeated any number of times, as is rotation. Both require control over physical transformations. In the case of superposition the position of the layers must be controlled precisely, and in the case of rotation, the angle must be controlled precisely. Are there such functions in the brain?

The process required in the brain is analog. For over 25 years, a dispute has raged over whether brains support analog processes or whether mental imagery is driven by non-analog means (Pylyshyn 2001). We can sidestep this question, though, by appealing to an in principle distinction between types of processes. In an important paper, Von Neumann (1948) mentioned that some processes in nature might be irreducibly complex. Any description of one of those processes would be as complex as the process itself. Thus, to simulate or model that process one would have to recreate all the factors involved. This holds whether the simulation or modeling is being performed internally or externally. Von Neumann put it like this:

> It is not at all certain that in this domain a real object might not constitute the simplest description of itself, that is, any attempt to describe it by the usual literary or formal-logical method may lead to something less manageable and more involved (p. 311)

Marr (1977) invoking the same idea, spoke of Type 2 processes where any abstraction would be unreliable because the process being described evolves as the result of "the simultaneous action of a considerable number of processes, whose interaction is its own simplest description" (p. 38). Protein folding and unfolding are examples of such processes, according to Marr.[7] Other examples might be the n

---

[7] From Marr (1977, p. 38) "One promising candidate for a Type 2 theory is the problem of predicting how a protein will fold. A large number of influences act on a large polypeptide chain as it flaps and flails in a medium. At each moment only a few of the possible interactions will be important, but the importance of those few is decisive. Attempts to construct a simplified theory must ignore some interactions; but if most interactions are crucial at some stage during the folding, a simplified theory will prove inadequate. Interestingly, the most promising studies of protein

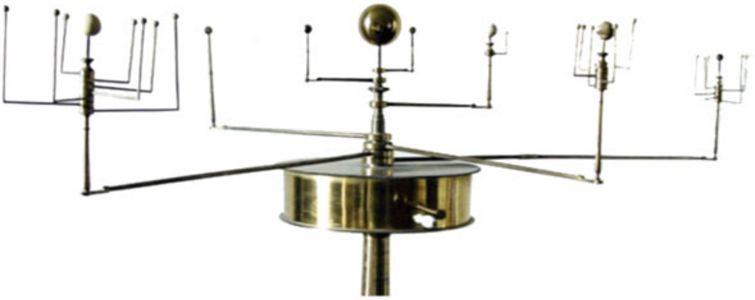


Fig. 4.11 In this mechanical orrery by Gilkerson, housed in the Armagh Observatory the movement of the planets and their moons are mechanically simulated. It is not possible to access an arbitrary position of the system without moving through intermediate states. This is a feature of simulation systems: they do not have a closed form or analytic solution. To compute the state of the system at t_{12} one must determine the state at t_{11} and move from there

body problem, the solution to certain market equilibrium problems, situations where the outcome depends on the voting of n participants, and certain quantum computations.

The hallmark of these problems is that there exists physical processes that start and end in an interpretable state, but the way they get there is unpredictable; the factors mediating the start and end state are large in number, and on any individual run are impossible to predict. To determine the outcome, therefore, it is necessary to run the process, and it is best to run the process repeatedly. No tractable equation will work as well.

How are these problems to be solved if we have no access to the process or system itself? The next best thing is to run a physically similar process. For example, to compute the behavior of an n body system, such as our solar system, our best hope is to construct a small analog version of that system—an orrery—then run the model, and read off the result (see Fig. 4.11). Using this analog process, we can compute a function (to a reasonable degree of approximation) that we have no other reliable way of computing.

The implication is that for brains to solve these sort of problems, it would be necessary for them to encode the initial state of the type II system, and then simulate the physical interaction of its parts. If this interaction is essentially physical—if for instance, it relies on physical equilibria, or mechanical compliance, or friction—there may be no reliable way of running an internal simulation. We need the cognitive amplification that exploiting physical models provides. We would need to

---

(Footnote 7 continued)

folding are currently those that take. a brute force approach, setting up a rather detailed model of the amino acids, the geometry associated with their sequence, hydrophobic interactions with the circumambient fluid, random thermal perturbations etc., and letting the whole set of processes run until a stable configuration is achieved (Levitt and Warshel 1975).”

rely on the parallel processing, the physical interaction, and the intrinsic unpredictability of those analog systems.

The conclusion I draw is that to formulate certain thoughts and to transition to others, we must either be able to represent arbitrarily complex states—states that cannot be represented in compact form—or we must rely on the external states themselves to encode their values and then use them to transition to later states. These external states we are able to name but never characterize in full structural detail.[8]

Conclusion

In order to extract meaning, draw conclusions, and deepen our understanding of representations and the world more generally, we often mark, annotate and create representations; we rearrange them, build on them, recast them; we compare them, and perform sundry other manipulations. Why bother? Minds are powerful devices for projecting structure on the world and imagining structure when it is not present. Our inner mental life is plastic and controllable, filled with images of speech, visual scene, and imageless propositions. For most of intellectual history this impressive capacity has been assumed sufficient for thought. Why do we bother to interact so much?

I have argued that much of thinking centers on interacting with external representations, and that sometimes these interactions are irreducible to processes that can be simulated, created, and controlled in the head. Often, the reason we interact with external representations, though, boils down to cost. Nothing comes without a cost. A useful approach to understanding epistemic interaction is to see it as a means of reducing the cost of projecting structure onto the world. To solve a geometric problem we might imagine a structure and reason about it internally, we might work with an illustration and project extensions and possibilities. At some point though the cost of projection becomes prohibitive. By creating external structure that anchors and visually encodes our projections, we can push further, compute more efficiently, and create forms that allow us to share thought. I have presented a few of the powerful consequences of interaction. It is part of a more general strategy that humans have evolved to project and materialize meaningful structure.

---

[8] In practice, though not in principle, computers fall into this category. When a workplace has been augmented with tools such as wizards, software agents and the like, it is possible to multiply the potency of basic strategies of interaction to the point where such increases qualitatively change what humans can do, what they can make sense of, and so on. Sometimes our best tools are analog, however, and these are the ones that may provide in principle augmentations to human thought.

References

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4,* 417–423.

Chabris, C. F., & Hearst, E. S. (2003). Visualization, pattern recognition, and forward search: Effects of playing speed and sight of the position on grandmaster chess errors. *Cognitive Science, 27,* 637–648.

Chambers, D., & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance, 11,* 317–328.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford: Oxford University Press.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences, 8,* 280–285.

Hutchins, E. (2001). Cognitive artifacts. In R. Wilson & F. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences*. Cambridge, MA: MIT Press.

Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics, 37,* 1555–1577.

Forsythe, W. (2008). http://www.youtube.com/watch?v=0P_4D8c2oGs&feature=related.

Kirsh, D. (1992). *When is information explicitly represented? The Vancouver Studies in Cognitive Science* (pp. 340–365). Oxford: Oxford University Press.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence, 73,* 31–68.

Kirsh, D. (1996). Adapting the environment instead of oneself. *Adaptive Behavior, 4,* 415–452.

Kirsh, D. (2009a). Problem solving and situated cognition. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 264–306). Cambridge: Cambridge University Press.

Kirsh, D. (2009b). Projection, problem space and anchoring. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2310–2315). Austin, TX: Cognitive Science Society.

Kirsh, D. (2009c). Interaction, external representation and sensemaking. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1103–1108). Austin, TX: Cognitive Science Society.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic actions. *Cognitive Science, 18,* 513–549.

Levitt, M., & Warshel, A. (1975). Computer simulation of protein folding. *Nature, 253,* 694–698.

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove, UK: Lawrence Erlbaum Associates.

Marr, D. (1977). Artificial intelligence: A personal view.". *Artificial Intelligence, 9,* 37–48.

Nagel, E., & Newman, J. (1958). *Gödel's proof.* New York: New York University Press.

Pylyshyn, Z. (2001). Is the imagery debate over? If so, what was it about? In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 59–83). Cambridge, MA: MIT Press.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Sacks, O. (2008). *Musicophilia: Tales of music and the brain*. New York: Alfred A. Knopf.

Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. London: Methuen.

Von Neumann, J. (1948/1961). General and logical theory of automata. In A. H. Taub (Ed.) *John von Neumann: Collected works. Volume V: Design of computers, theory of automata and numerical analysis* (pp. 288–326). Oxford: Pergamon Press.

Vygotsky, L. (1986). *Thought and language* (rev ed.). Cambridge, MA: MIT Press.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Basil Blackwell.

Chapter 5
Human Interactivity: Problem-Finding, Problem-Solving, and Verbal Patterns in the Wild

Sune Vork Steffensen

Abstract This chapter presents an interactivity-based approach to human problem-solving in the wild. It introduces the notion of 'interactivity', here defined as sense-saturated coordination that contributes to human action. On this view, interactivity is an ontological substrate that can be studied as interaction, as cognition, or as ecological niche production (section "Human Interactivity"). While the chapter theoretically argues in favour of a unified, anti-disciplinary approach to interactivity, it turns its attention to the cognitive ecology of human problem-solving. It does so by presenting a method of Cognitive Event Analysis (section "From Interactivity to Cognitive Events"), that leads to a detailed analysis of how a problem in the wild is being solved (section "The Invoice Case"). The analysis addresses the cognitive dynamics of how two persons in a work setting reach an insight into the nature of a problem, including the spatial organisation of the workplace, the interbodily dynamics between the two participants (especially in relation to gaze and the manual handling of papers), and verbal patterns that prompts them to simulate how the problem appears to a third party. The chapter concludes that human problem-solving is far less linear and planned than assumed in much work on the topic. Rather than problem-solving, it appears as solution-probing in real-time. The cognitive trajectory to a viable solution is thus self-organised, unplanned, and on the edge of chaos.

Introduction

The biggest problem in finding a needle in a haystack is that in most haystacks there are no needles. Comparably, in problem-solving the real challenge is problem-finding, that is identifying the nature of the problem. In problem-finding, our everyday experience differs immensely from what is being examined in the

---

S.V. Steffensen (✉)
Department of Language and Communication, University of Southern Denmark,
Campusvej 55, 5230 Odense M, Denmark
e-mail: s.v.steffensen@sdu.dk

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_5

laboratory settings that nurture cognitive psychology. In this artificial setting subjects are typically presented to a well-defined problem; for example “connect the nine dots with four connected straight lines without lifting your pencil from the paper” (Weisberg and Alba 1981, p. 170; the nine dot problem was originally presented in Maier 1930). In contrast, real-life existence does not provide us with the luxury of pre-established, well-defined problems. We rarely meet instructions like “go down to the diner on the corner, order a cheese burger, bring it back to your office and eat it.” We just feel hunger. From that feeling, we need to figure out what to do. Human life unfolds in an indefinite problem space. Hutchins’s (1995) is correct: *Cognition in the wild* is very different from the laboratory activities that cognitive psychologists for a century have investigated as problem-solving. However, as this chapter demonstrates, in most real-life contexts, *cognition in the wild* is even wilder than generally assumed in the Distributed Cognition literature. Rather, it is profoundly interpenetrated with the real-time flow of human co-existence in a distinct human ecology: that is, how human beings engage in collective and socio-culturally enhanced problem-searching, problem-finding and problem-solving activities in a self-organised problem space. Following Kirsh (1997), Steffensen (2012) and Pedersen (2012, 2015), I refer to this real-time flow of co-existence as *interactivity*. While the concept is clarified below, interactivity is defined as *sense-saturated coordination that contributes to human action*.

As interactivity comprises all aspects of this flow of co-existence, it has a primordial quality that does not fit readily into how science is compartmentalised. Thus, nowhere has the luxury of predefined problem spaces been more emphasised than where disciplines each seek to grasp their small part of reality.[1] Thus, what the cognitivist perceives as problem-solving, the microsociologist interprets as social interaction, and the biologist construes as human ecological niche construction. Underlying each perspective, one finds interactivity: it is an ontological substrate that each discipline has turned into an ‘object’. While all three perspectives may yield descriptively adequate models within an epistemological domain (cf. Hemmingsen 2013), they cannot, in themselves, provide an explanatory model of interactivity, of what really happens in the flow of human existence. Though we should not delude ourselves by believing that we are even close to such a model, a focus on interactivity may provide a grounding for a theoretical and methodological approach to what Andy Clark has called *The Heideggerian Theater*: “Our target is not just a neural control system but a complex cognitive economy spanning brain, body and world. Within this complex economy, the body plays a crucial role. […] The body is—dare I say?—the Heideggerian Theater: the place where it all comes together, or as together as it comes at all” (Clark 2008, p. 217).

In contrast to Clark, and many recent approaches in cognitive science, a focus on interactivity rejects organism-centrism. It does not pivot on the living *body* but,

---

[1] Saussure was indeed correct when he, a century ago, established that “c’est le point de vue qui crée l’objet” (Saussure 1916/1973, p. 23).

rather, on *bodies-in-action*. Accordingly, rather than appeal to the utility functions of a cognitive *economy*, the emphasis turns to how we act viably in a cognitive *ecology*. I thus maintain that this approach is better equipped for actually responding to Clark's desideratum or the aspiration to rethink human nature as "not accidentally but profoundly and continuously informed by our existence as physically embodied, and as socially and technologically embedded, organisms" (Clark 2008, p. 217).[2]

This chapter begins with a theoretical discussion (section "Human Interactivity") of interactivity, follows up with a methodological intermezzo (section "From Interactivity to Cognitive Events") and then presents a case study of human interactivity (section "The Invoice Case"). In section "Conclusion: Interactivity, Language and Cognition", I draw some conclusions on how the approach can contribute to the understanding of language and cognition.

Human Interactivity

Defining interactivity as sense-saturated coordination that contributes to human action characterises three aspects of the relevant phenomena. First, *coordination* refers to a reciprocal flow of minuscule, pico-scale interbodily movements that link and lock human beings in self-organised systems. This basic insight is shared by work that invokes concepts such as *distributed cognitive systems* (Hollan et al. 2000), *situated activity systems* (Goodwin 2000), and dialogical systems (Steffensen 2012). Second, this coordination is *sense-saturated*, that is, it is pervaded by our species-specific capability for sense-making (Linell 2009; Steffensen 2015). We engage in sense-making as our bodies integrate present circumstances with autobiographic memories and sociocultural histories: through sense-making the not-here and the not-now saturate our here-and-now coordination. Third, sense-saturated coordination constrains what we do and how we do it. For instance, if you grew up with the habit of greeting through cheek kissing, you know on which cheek to start, and you know how many kisses to exchange. But if you are not accustomed to the habit, and you greet a person who is, you can still engage in the social practice of greeting-through-cheek-kissing, but you are to a wide extent dependent on following the dynamics of the other person's body. This example illustrates the power of the interbodily dynamics, our interbodily agility as we mutually interpret and anticipate each other's movements, and the constraining and enabling function of socioculture. Because interactivity is sense-saturated, our actions and experiences are, at once, situated and non-situated, and we are furthermore bound to overthrow the monolithic distinction between the realm of sense-making (e.g., 'language') on the one side and the realm of behaviour on the other (cf. Steffensen 2011). Thus,

---

[2] For a fuller critique of Clark's position, see Steffensen (2009, 2011).

what we normally conceive of as ‘human language’ is a pattern in interactivity, and as such it is always dynamics, as well as symbolic (Cowley 2011).

The interplay between interbodily dynamics and sociocultural constraints on behaviour prompts the approach to adopt a systemic focus on *results*, “not as a simple effect or consequence of behavior but [as] a new possibility of action forming in the process of transition from one act to another” (Järvilehto 2009, p. 116). In focusing on results, the Heraclitean flow of interactivity ceases to be either a sequential string of separate stimuli and responses or an incessant flow of undifferentiated, unstructured “pure being.”

The concept of ‘human interactivity’ grounds anti-disciplinary (*pace* Peter Jones) empirical work that attempts to move beyond the microsocial study of social interaction, beyond the cognitive study of functional, computational systems and minds, and beyond the study of biological organisms *sui generis*. It presupposes a crucial distinction between ‘interaction’ and ‘interactivity’; whereas interaction captures a relation of dependence between separable systems, interactivity explores their bidirectional coupling. This view contrasts with a focus that seeks out the natural laws that describe, for example, how the earth and the moon interact (cf. Kirsh 1997, p. 83). But it also contrasts with approaches that invoke social rules in order to explain how separate human agents orient to each other and the world (cf. Hodges and Baron 1992). As social rules are violable, human interaction becomes purely normative. In ethnomethodology and conversation analysis, therefore, human interaction is a *contingent* co-construction of negotiated meaning. Interestingly, both celestial (i.e., natural laws) and social interaction (i.e., social rules) can be studied by strict application of inductive methods. In both domains, theorists provide formulations or *models* of how inert bodies or social actors affect each other (cf. Sacks et al. 1974). And just as astronomers need not claim any larger purpose in nature, interaction analysts avoid issues of intentionality. They neither allow biological bodies to *become* social actors or ask *why* people act as they do. Any method based on studying interactional regularities and repeated patterns will fail to capture the biological dynamics of human interaction or its (cognitive) *results*. They simply overlook variable aspects of the interactivity that connects living beings.

Interactivity takes us beyond the computational functionalism of cognitive science (including that of classical Distributed Cognition). Functionalists propose that computational processes can take place both inside and outside living beings, irrespective of the medium in which they unfold. By contrast, the concept of ‘interactivity’ views the “glue of cognition” (Kirsh 2006, p. 250) as “complex, dynamic coupling between two or more intelligent parties” (Kirsh 1997, p. 83). It thus makes a difference how a cognitive system balances animate agency with non-animate contributions. Thus, if the cognitive work is distributed between animate parties (human beings, to simplify the picture), “the involved parties must *co-ordinate* their activity or else the process collapses into chaos; all parties exercise *power* over each other, influencing what the other will do, and usually there is some degree of (tacit) *negotiation* over who will do what, when and how" (Kirsh 1997, pp. 82–83; emphasis in the original). Cognition depends on the total organisation of organisms' self-reflexive being in their shared environment. On this view, human *interactivity* is a non-local phenomenon (Steffensen and Cowley 2010; Steffensen 2015): it links the biological organism with autobiographical memory, socio-cultural resources (e.g., verbal patterns) and environmental structures. Interactivity is thus whole-bodied activity that flows between human beings—or between a single human being and cultural artefacts and procedures (e.g., previously crafted texts or technological devices). To generate *joint results*, it is enough that we coordinate and strive to make things happen.[3]

Along with other contemporary work in cognitive science, this approach challenges biological reductionism. Living organisms perceive *as* they act and, conversely, act *as* they perceive; far from being an inner process, cognition exploits the full organisation of embodied and emotional action-perception cycles (Kirsh 1997; Järvilehto 1998; Robbins and Aydede 2009). This view thus traces cognition to a history of how living systems adapt within a changing environment. Situated, living, cognitive systems thus function across the boundary of the skin; accordingly, we take an *ecological* stance that overthrows the dichotomy of computational reductionism and biological reductionism. While the former reduces cognition to its functional properties, overlooking feelings and the properties of bodies, the latter stick to dermal metaphysics: on a priori grounds, the skin is taken to define the boundary of what counts as cognitive. In contrast, a systemic approach to interactivity pivots on an extended ecology (Steffensen 2011) that defines human cognition in terms of the flexibility and adaptivity of (human) organism-environment-systems (Järvilehto 1998, 2009).

A focus on interactivity obliterates any sharp boundary between the biological and the social. First, social normativity informs and constrains our biological being, for example by filtering out activities that are judged inappropriate in certain social contexts (e.g., picking one's nose). Second, and crucially, our very existence depends on conspecifics and the extended ecology. Human development depends on caregivers (e.g., Trevarthen 1998; Bråten 2009) who enable the new-born infant to learn from being tightly coupled to human beings in its environment. Plainly it is unwarranted to limit biology to the body: as we breathe, move, touch, act, perceive and care, we *radiate* into our surroundings. Our being is interbodily being.

---

[3] In line with this view, Donald (2001) argues that humans alone developed a cultural capacity to *voluntarily* retrieve experience. Culture turns our social and physical environment into a *cognitive* resource, and not just an independent, external material resource. Therefore, studying cognition as a Distributed Cognitive System requires that we take into account longer time-scales than that of here-and-now situated interactions, for instance those involved in autobiographical memory (Donald 1991, 2001, 2012; cf. Cowley 2012), collective memory (Wertsch 2002) and cultural artefacts.

From Interactivity to Cognitive Events

Interactivity is a primordial substrate of human life. Phenomenologically, we may perceive it as ‘language’, ‘interaction’, ‘cognition’ or ‘nice-construction’, but these are only perspectives on interactivity, not ontologically real phenomena per se. Having made this claim, I now turn to how it serves human beings as they find and solve problems in the wild, that is, in the real-time flow of day-to-day existence. Accordingly I focus on the *cognitive* dynamics of interactivity that take part in what, following Hollan et al. (2000), I call *Distributed Cognitive Systems* (henceforth DCS). The DCS is a self-organising entity that arises as human beings co-engage through interactivity, and connect up brains, bodies and aspects of the environment. The dynamics of the DCS derive from how one or more people engage with each other and external artefacts. Such a distributed system has emergent cognitive properties that differ radically from those of the components; moreover, these properties cannot be inferred from the properties of the components, no matter how much we know about the details of the properties of those components (cf. Hutchins 1995).

The cognitive emphasis also reflects in the term for the method applied in this chapter, namely *Cognitive Event Analysis* which builds on Steffensen et al. (2010, 2016), Steffensen (2012), Galosia et al. (2010), Pedersen (2010, 2012). The method of analysis proceeds in two stages: first, since interactivity lacks inherent beginnings or endings, empirical happenings are defined by external criteria. The first stage is therefore an *event identification* stage. For instance, in a medical context the presentation of a given patient symptom can function as such a criterion, irrespective of how the involved participants orient to it, and irrespective of whether it is noticed or not. The second stage tracks the dynamics involved in the interactivity by means of an *event trajectory analysis*. The two stages have been extended into a five stages procedure in the treatment in Steffensen et al. (2016). For now, I stick to the two phases which will be elaborated below.

Event Identification

Given its primordial status, the current approach defines interactivity neither in terms of what ‘we’ do, say, or mean. Rather, it asks: *what happens*? Galosia et al. (2010) uses this question to define events, that is “*changes in the layout of affordances*” (Chemero 2000, p. 39; emphasis in the original) which yield *results*. This approach is indebted to Timo Järvilehto’s *systemic psychology* (or *organism-environment systems theory*), according to which “the research should start from the determination of the results of behavior and lead to the necessary constituents of the living system determining the achievement of these results” (Järvilehto 2009, p. 118). Results are typically, though not necessarily, triggered by how people draw on sense-saturated coordination; they are, however, not necessarily related to a participant’s desired outcomes.

Given a set of criteria that are relevant to the events (e.g., diagnosing a patient), the results are *identifiable*. However, there is no reason to think that interactivity depends on a determinate set of results—what happens will not be the same for different participants and investigators. By tying results to external criteria, they become a probe that serves to explore the cognitive ecology of human interactivity. The focus on results obviously links the approach to cognitive science, and indeed traditional and, by extension, lab-based cognitive psychology plays an important part when it comes to formulating external criteria that define results. From an experimental perspective, problem-solving can be seen as 'search' in a temporally organised problem space, where the lab subjects are requested to move from a 'before' to an 'after'. In real life, as illustrated below, this temporal dimension often appears post hoc. In other words, identifying temporal patterns that resemble a journey through a problem space give the event a pseudo-task structure. A results-based approach to interactivity thus offers criteria for zooming in on interactivity to clarify how salient events unfold at moments of an unending process. In other words, this approach provides a method for identifying *how* cognitive events arise from the flow of human interactivity.

Event Trajectory Analysis

An *event trajectory* is the path taken by a DCS as it moves through an infinite problem space in a way that yields results. Assuming that distributed cognitive systems are *animated* by persons, such an analysis focuses on the dynamics of what people *do*, as they perform various tasks. Thus a cognitive event analysis clarifies how the DSC reconfigures to yield results and how these are generated. Reconfigurations and results act as *transition points* on the event trajectory: the pattern and timing of these transition points thus give each event trajectory a unique 'fingerprint'. In events where problems are solved, the emergence of the solution defines a transition point that divides the event trajectory into a 'before' (big problem, no solution) and 'after' (good solution, no problem). Such a transition point is more salient than other transition points: it constitutes the *event pivot* or cognitive origo. Its importance lies in how it alters the cognitive trajectory: *before* the event pivot, the participants seek a solution to the problem; *after* the event pivot they react to what they have found. As the event pivot constitutes a cognitive origo, Cognitive Event Analysis (predominantly, but not exclusively) uses the convention of timing the event trajectory in relation to the event pivot and in milliseconds (ms). Other important transition points in the event trajectory include (a) the first observation of the problem; and (b) how unsuccessful strategies are gradually replaced by more successful ones. The case study below gives a more thorough introduction to the event trajectory, event transition points, and the event pivot.

Components in Distributed Cognitive System

Since interactivity flows across human agents and non-human artefacts, it becomes vital to establish what does and does not, for a given event, function as parts of the cognitive system. Following Clark's (2008) *Principle of Ecological Assembly*, in which it is assumed that "the canny cognizer tends to recruit, on the spot, whatever mix of problem-solving resources [that] will yield an acceptable result with a minimum of effort" (p. 13), I define the reconfiguration of a DCS in terms of how it includes and excludes neural, bodily, worldly, virtual or historical structures. This description of the DCS as a plastic self-reconfiguring system resembles Wilson and Clark's (2009: 65) description of a "transient extended cognitive system," which they define as "a soft-assembled whole that meshes the problem-solving contributions of the human brain and central nervous system with those of the (rest of the) body and various elements of local cognitive scaffolding." The main differences between the position advocated here and that of Wilson and Clark, is, first, that their focus is on how the individual cognizer ("the human brain" in the singular) recruits bodily and environmental structures in the cognitive meshwork, while I focus on how the interpersonal dynamics in a dialogical system constrain the cognitive trajectory of the whole system (cf. Steffensen 2011). Second, the ecological-observational study presented here explores *how* such a system manages its self-reconfiguration, while Wilson and Clark seem to be more interested in the contributions of various (biological and non-biological) parts to the extension of the human mind. This, however, does not clarify *how* the plasticity of cognitive systems is achieved.

In spite of these differences, both positions seem to agree that there no a priori basis for specifying what is part of—or can become part of—a cognitive system. Thus, when a DCS is said to consist of X, Y and Z structures, it is the result of a post hoc procedure where the temporal event trajectory is translated into a static cognitive inventory. The circularity of this procedure is non-trivial: while an event trajectory is non-linear and unpredictable, the component structures that are (or can become) part of the DCS constitute affordances that *constrain* the infinite problem space of the DCS. This, indeed, is why the methodology presented here uses a heuristic in seeking out how solutions are reached. This heuristic consists of a tentative hierarchy of component structures that are likely to contribute to the DCS. First of all, as agreed on by both the TECS and the DCS approach, a cognitive system is animated by human beings, and therefore it requires human components (for at least some of its trajectory). Accordingly, persons are regarded as an especially salient part of a DCS. However, within that category, many heterarchical relations can be involved; for instance, one person may be organisationally responsible for directing the DCS, or two participants can take turns as main cognizer (cf. Galosia et al. 2010). Second, especially in the "highly rationalized environments" (McGarry 2005, p. 187) described in the distributed cognition literature, *technologies* that were specifically designed to serve in a DCS are likely to play a major role (at some stage at the event trajectory, at least). This applies, for example, in Hutchins's (1995b, p. 274) description of the airspeed indicator in a McDonnell Douglas MD-80. However, this cannot be generalized as a rule: an alarm system will do no good in the DCS if a human component has turned it off or fails to pay attention. Third, many kinds of *artefacts* are redeployed in the DCS, either because of their material affordances (cf. the painter who uses his brush to measure a scenery), or because human beings interpret them as giving off information. Perry (this volume) provides a good example: "if someone had a pair of muddy boots under their desk, it meant that they have been on the [construction] site and could be asked about the current work situation" (Perry this vol.). Fourth, I stress Hollan et al.'s (2000, p. 176) insight that cognition "may be distributed through time in such a way that the products of earlier events can transform the nature of later events." In workplace settings, such temporally distributed parts of the DCS often appear as *procedures* that, like second order language (cf. Thibault 2011), constrain human behaviour. Procedures aim at standardising the performance, typically through written instructions, but these may also be oral or memorised. The inclusion of memorised procedures in this hierarchy rests on an invert version of Clark's parity principle (Clark and Chambers 1998): if a standardisation counts as a procedure in its written form, we also accept memorised aids that afford uniformity in performance as procedure. Fifth, we can also include *narratives*. Narratives provide folk-theoretical schemes for understanding causality, ethics, emotions, for example. As such, narratives provide short term and long term constraints on behaviour.

The Invoice Case

The following data were collected by Anne-Kathrine Brorsen during her investigation of employer-employee interaction in a Danish company (Brorsen 2010). The data consist of 12 h of video recordings and, in most of the data, the video recorder serves as a "silent observer" (Harel 1991). What is observed here is the workplace of one person sitting on a chair at his desk with a computer in front of him. From time to time, a co-worker appears and stands next to the first worker, while they interact about various issues. Both men are dressed in jeans and t-shirts, and I refer to them by the colours of their shirts. Thus, the man sitting at his work-station is *Black*, while his co-worker is *White* (see Fig. 5.1 for an overview of the lay-out).

Cognitive event analysis was used to identify a single instance of problem solving from the massive data set. It is thus a true instance of problem-solving in the wild: the two parties, led by Black, try to ensure that the company's invoices have the required Company Identification Number or '*CVR number*' (a CVR number is a code that the state agency *Det Centrale Virksomhedsregister (CVR,* The Danish Central Business Register) supplies to all Danish companies. If this number does not appear on an invoice, it cannot be paid by the clients invoiced. In this situation, having collected a document from printer, no CVR number appears on it. This is the

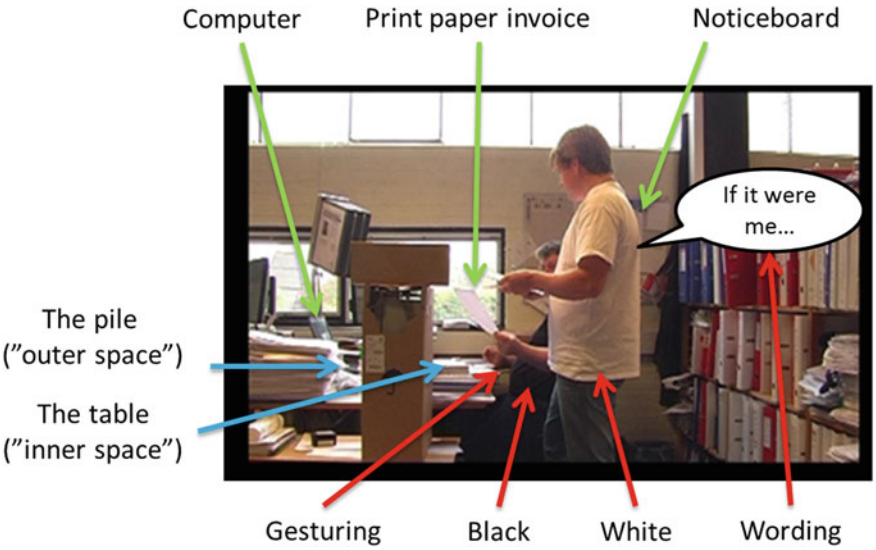
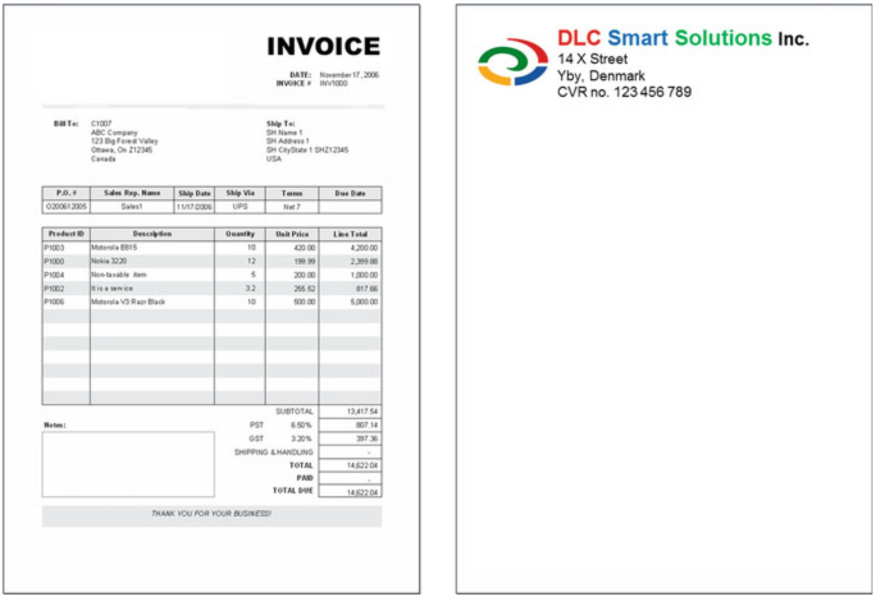


Fig. 5.1 The distributed cognitive system in the invoice case. The *red arrows* point to the human components of the DCS; the *green arrows* to the artefactual components; and the *blue arrows* to the spatial organisation of the DCS's environment

problem that Black and White faces as illustrated in Fig. 5.2. The left half of the figure reconstructs the printed invoice that the two men see; no CVR number appears; the right half is an anonymised reconstruction of the company's logo paper. As shown, the CVR number is in fact pre-printed under the company's name and address. Since the company would always use logo-headed paper when sending an invoice to a third party, there should be *no problem*. However, in preparing the invoice Black and While have printed it on blank paper. Unbeknown to them, the problem arises from the choice of tray in the office printer! In principle, the *solution* is simple: they have to realize that they should discard the draft invoice for the version that their customers would receive. But as we all know from everyday life, problems are only simple when one knows the solution.

In this case, the salient cognitive event is identifying the problem: the event pivot arises in a 1900 ms sequence where White clearly articulates the solution: *nå nej men det er der jo hvis vi printer ud på logo papir* (Eng. 'well no but it [the number] is there if we print on logo paper'). According to the time notation convention the event pivot starts at 0 and ends at +1900. A second relevant event trajectory transition point arises as the participants shift from being unable to solve the problem, to reframing it in a way that gives a solution. This transition point occurs at −8575. Given these transition points, the event trajectory runs through the three phases shown in Fig. 5.3.

The first phase is characterised by unsuccessful attempts to reach a solution. This phase ends at the problem reframing point. The second phase occurs between the



INVOICE

DATE: November 17, 2006
INVOICE # INV1000

Bill To: C1007
ABC Company
123 Big Forest Valley
Ottawa, On Z12345
Canada

Ship To:
SH Name 1
SH Address 1
SH CityState 1 SHZ12345
USA

P.O. #	Sales Rep. Name	Ship Date	Ship Via	Terms	Due Date
0200612005	Sales1	11/17/2006	UPS	Net 7	

Product ID	Description	Quantity	Unit Price	Line Total
P1003	Motorola E815	10	420.00	4,200.00
P1000	Nokia 3220	12	199.99	2,399.88
P1004	Non-taxable item	5	200.00	1,000.00
P1002	It is a service	3.2	255.52	817.66
P1006	Motorola V3 Razr Black	10	500.00	5,000.00

Notes:

SUBTOTAL		13,417.54
PST	6.50%	807.14
GST	3.20%	397.36
SHIPPING & HANDLING		-
TOTAL		14,622.04
PAID		-
TOTAL DUE		14,622.04

THANK YOU FOR YOUR BUSINESS!

DLC Smart Solutions Inc.
14 X Street
Yby, Denmark
CVR no. 123 456 789

Fig. 5.2 The cognitive problem and its solution. *To the left* the figure shows a reconstruction of the print paper invoice that the two men are handling. *To the right* it shows a reconstruction of the company's logo paper with the CVR number under the company's address

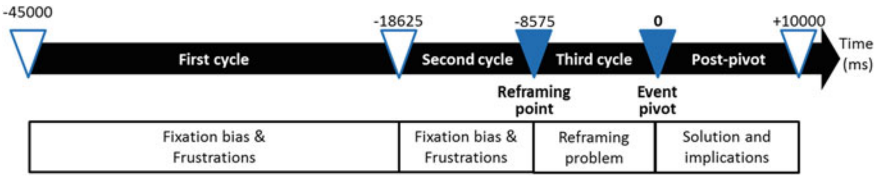


Fig. 5.3 The event trajectory. The figure shows the timeline of the event trajectory with indications of cycles and phases; *blue triangles* mark the two main transition points, the problem reframing point and the event pivot; *white triangles mark* secondary transition points. The *boxes* indicate the main character of the phase/cycle in question

trajectory transition points as the problem is reframed; finally, in the third phase, a solution emerges after the event pivot. Below I analyse the event trajectory in these three phases while giving emphasis to the cognitively salient problem reframing that occurs between the two event pivots. The purpose of the analysis is to demonstrate that problem-finding, problem-solving and cognitive events do not just appear in an *aha*-like fashion: they depend on the flow of interactivity, as persons immethodically zigzag along a cognitive trajectory in an indefinite problem space, in their search for something that can work as a problem solution.

Frustration, Fixation and the Elusive Problem

As we enter the situation at −45000, Black and White are blocked by an ill-defined problem that seems to defy solution. Black repetitively restates *what* the perceived problem is. The pattern recurs three times, giving rise to three problem-solving cycles in the pre-pivot phases shown in Fig. 5.3. The first two cycles run from −45000 to −18625 (ca. 26 s) and from −18625 to −8575 (ca. 10 s), respectively, while the third links the two event pivots, from −8575 to 0 (ca. 8.5 s). As in Steffensen et al. (2010), the presentation of these cycles takes a starting point in the micro-scale of the verbal patterns:

1. B: men (.) jeg kan fortælle dig der er ikke nogen som helst der
2. vil betale den faktura der
3. W: nej det er jeg da godt klar over. [...]
1. B: but (.) I can tell you there is no-one whatsoever who will
2. pay that invoice there
3. W: no I am aware of that. [...]

As shown by White’s response in (3), he is already aware of the nature of the problem. At this moment, the information bearing properties of language—its symbolic function—is negligible. Often, it is simply false that language acts as “one of the structured representations produced and coordinated in the performance of the task” (Hutchins 1995, p. 231). At best, it functions as a means of concerting their attention; at worst it takes them round in circles, preventing them from identifying the problem. Failure to come up with a solution frustrates the two participants as shown by two manifest outcomes. First, White acts out his frustration by blaming a party (unknown to the observer) that is referred to as *dem* (English: ‘them’).

3. W: [...] Det er derfor vi sagde det
4. til dem at den ikke duede jo. (0.7) Men det skulle vi ikke
5. blande os i fordi det var som det var aftalt. (1.3)
6. så var det det.
7. B: står der et CVR nummer hernede et sted?
8. W: = men det kan være du kan få et bedre svar end jeg ku,
9. jeg fik bare at vide at sådan er det.
10. [...]
3. W: [...]. That’s why we told them
4. that it was no good. (0.7) But that was not our business
5. because it was as agreed upon. (1.3)
6. So that was it.

```
7. B: is there a CVR number down here somewhere?
8. W: = but it might be that you can get a better answer than I
9. could, I was just told that that was how it was
10. […]
```

Even weak versions of the so-called Frustration-Aggression Hypothesis (Dollard et al. 1939; Berkowitz 2007) would see this as predictable. Frustrations, it is found, tend to instigate aggressive or reproachful feelings by attributing emotion factors that lie beyond individual control. Second, and productively, the frustration instigates attempts at reframing. Thus, as White pauses in line 4, Black turns to his noticeboard on the right; then, 5400 ms after turning, he explains what he is looking for (line 7): *står der et CVR nummer hernede et sted?* ('is there a CVR number down here somewhere?'). Black's engagement in the interactivity is driven by emotions, frustrations, and the available resources at hand. All in all, Black spends about 20 s looking for the CVR number on the noticeboard. This is hardly surprising: in a cognitive ecology, participants look for solutions where they, based on previous experience, think the problem is. At least since Scheerer (1963), this phenomenon has been associated with *fixation bias* and, as well known it rarely leads to successful outcome. Presumably, Black is looking for a CVR number that can be fed into the computer as part of the printed section of the invoice. As Weisberg and Alba (1981, p. 188) note, "one of the difficult aspects of these problems may be that it is not clear that the obvious solutions will not work," and indeed Black's 26 s of problem probing cycle did not work. Immediately after the cycle, he finds himself back at the initial point in the problem space. A common sense approach to problem solving suggests that if a problem solving strategy fails, another one should be attempted. However, this does not happen: Black embarks on his second cycle of the same strategy, and the repetitive strategy leads to recycling the verbal pattern from line 1: *men jeg kan fortælle dig* (English: 'but I can tell you'), repeated in line 11. Once more, Black recapitulates the problem, and, once more, White does little more than agree (line 14):

```
11. B: men jeg kan fortælle dig, den her den lader de bare ligge.
12. Den her den betaler de aldrig nogensinde. (3.0)
13. Den vil aldrig nogensinde blive betalt, den her (1.0)
14. W: Nej nej
11. B: but I can tell you, this one they'll just discard.
12. This one they will never ever pay it. (3.0)
13. It will never ever get paid, this one. (1.0)
14. W: No no
```

However, in complex dynamical systems, there is no such thing as replication of pattern. While the wording may be identical, its physical features and the context of situation can be used by the dialogical system (cf. Steffensen 2012). As shown in Fig. 5.3, the second cycle sets up *enabling conditions* for the transition point at −8575, that is, the problem reframing point. Although unsuccessful in its own terms, the second cycle brings about the interactivity that will later provide the desired solution.

Methodologically, this insight obviously uses a post hoc procedure where the salient parts of the third cycle in Fig. 5.3 are traced to the second cycle. So, what are the enabling conditions that lead to a successful third cycle? It will come as no surprise for scholars in distributed cognition that they depend on not only what goes on ‘in the head’ but also on the material artefacts. Thus, during the 2175 ms in line 11, Black reconfigures the DCS by means of *artefactual reconfiguration*. Using the invoice, which until this point in time has been lying on the table, Black picks it up and thus establishes an external focus of attention. This reconfiguration of the DCS exploits the interbodily dynamics of how Black manipulates the artefactual lay-out using his arms and hands. By so doing, Black prompts White’s gaze behaviour in a way that contributes to the reconfiguration the DCS (see Ball and Litchfield this vol.). This event is illustrated in Figs. 5.4 and 5.5.

As Black picks up the invoice at −18,150, his action is synchronized with uttering the demonstrative object deictic *den her* (literally ‘this here’; cf. Bang and Døør 2007). This visible movement prompts White to look at the invoice for 2250 ms. In so doing, they establish shared attention to the printed invoice that lies behind the fixation bias. Then, 1000 ms later, they establish eye contact for a remarkably long 1625 ms (from −14,125 to −12,500, corresponding to the first half of the pause in line 12). Eye contact apparently recalibrates the DCS around the interactivity: they face the same problem. Attention thus subtly shifts from a shared perceptual object to their role as perceptual *subjects*: each becomes aware of the other’s role in the problem-solving. Having established this interpersonal frame, Black turns his attention back to the invoice (at −11,800). As he does so, he uses a

Fig. 5.4 Black picks up the invoice. Picture 1 (−18150): Black reaches out and grasps the invoice. Picture 2 (−17025): Black holds the invoice and both participants gaze at it. Picture 3 (−14125): Black holds the invoice, and the two participants establish eye contact. Picture 4 (−10425): Black bobs the paper lightly

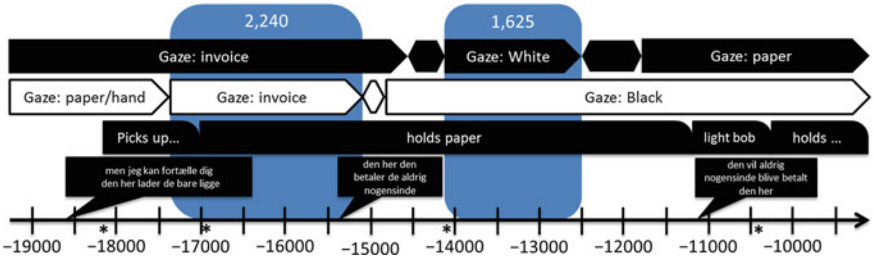


Fig. 5.5 The timeline of the second cycle (from −19000 to −9500). *Black fields* show Black's actions, *White fields* show White's actions. The *two top rows* show gaze direction (*pentagons*) interrupted by head turns (*hexagons*). The *third row* shows Black's hand/arm movements (*squares with a rounded corner*). The *fourth row* shows Blacks wording (in *squared speech bubbles*). The *blue areas mark* synchronicities in gaze: the shared attention to the invoice (2,240 ms) and the eye contact (1,625 ms). The *asterisks* at the timeline indicate the four points in time that correspond to the four pictures in Fig. 5.4

light hand movement to bob it up and down[4] while he utters the initial *den* ('it') in line 13: *Den vil aldrig nogensinde blive betalt, den her* ('It will never ever get paid, this one'). The two participants thus both attend to the invoice and to each other, while concerting the use of the linguistic deictics *den her* ('this one') and *den* ('it') with shifts in gaze and attention.

As part of this second cycle, the events bring about a subtle shift in deictic attention: whereas the first cycle centred on *de* ('they'), an unknown group in their own organisation, the same deictic now identifies the *receiver* of the invoice. In other words, the receiver *lader den ligge* ('just discards it') and thus *aldrig nogensinde betaler den* ('never ever pays it'). The way in which this subtle deictic shift (cf. Steffensen 2012) leads to the secondary event pivot (at −8575), is an example of what Lorenz (1963) famously described as when "slightly differing initial states can evolve into considerably different states," which later became the metaphorical butterfly who by flapping its wings in Brazil sets off a hurricane in Texas. In this case, the "butterfly" is the low-voiced, unnoticeable deictic entrance of the receiver, accompanied by a shift in attention, while the "hurricane" is the identification and, thus, solution, of the participants' real problem.

Re-framing the Problem

From an observer's perspective the problem is that the printed version of the invoice biases the participants' thinking. Accordingly, to reach a viable solution or result, they must dissolve their fixation bias. This dissolution occurs in the 8575 ms

---

[4] The light shake of the invoice is invisible in the still shots in Fig. 5.4, but a close examination of the 25 frames from the video recording reveals the movement.

between the trajectory transition points; it occurs between the problem reframing point and the event pivot. Further, as the terms suggest, they depend on reframing the problem. This event trajectory can be given a first approximation by considering a transcription of Black's speech:

```
15. B: Hvis det var mig så røg den bare hen i stakken. (0.7)
16. Den kan jeg ikke betale. (0.4)
17. Hvorfor kan jeg ikke det? (0.8)
18. Der er ikke noget CVR nummer på. (0.9)
19. Du må ikke sende en faktura uden CVR nummer (0.4)
15. B: If it were me then it just went in the pile (0.7)
16. I can't pay that. (0.4)
17. Why can't I pay it? (0.8)
18. There is no CVR number on it. (0.9)
19. You can't send an invoice without a CVR number. (0.4)
```

In section "Frustration, Fixation and the Elusive Problem", I argued that the second cycle of the event trajectory brings about the necessary conditions for reframing of the problem, namely: (1) the manipulation of the invoice; (2) the recalibration of the human parts of the DCS, through shared attention to each other and the invoice; and (3) the introduction of the invoice receiver, mediated by the personal deictic *de* ('they'), and the concurrent shifts in attention. Next, I focus on *how* these three elements dissolve the fixation bias. I argue that they are catalysed by a fourth and novel element which defines the secondary event pivot. The key to understand lies in the 1400 ms represented in line 15.

Once again, Black repeats that the problem is that the invoice cannot be paid. However, in this third recapitulation of the problem, he uses a remarkably different strategy; he does so by attributing a hypothetical *narrative* to the invoice receiver. Linguistically this is marked by a formula for hypothetical thinking: *hvis det var mig* (Eng. 'if it were me'). Cognitively, the subtle introduction of the receiver in line 11–12 affects the current trajectory by prompting Black to adopt an *alter-centric* perspective on the invoice. In so doing, he sees the problem from the *perspective of the invoice receiver*. In the tradition of dialogism, Linell (2009, p. 83) refers to this phenomenon as *alter-centric perception*, which he elaborates as follows: "The other's "outsideness" brings in a 'surplus' of vision, knowledge and understanding other than you had before or you had expected to encounter. The other may see things from points-of-view that have so far been strange or unfamiliar to yourself, and this forces you to reflect and try to understand, thereby possibly enriching your, and our collective, knowledge and language".

Linell's description focuses on situations where the two parts of this relation are co-present parties. However, alter-centric perception also reflects on non-present, third parties (Linell 2009; cf. Bang and Døør 2007). Indeed, this is why the narrative entails a recalibration of the deictic system, as Black invokes a first person deictic *jeg* ('I') to index the invoice receiver. Thus, whereas in the first cycle (line 12), Black said that *den betaler <u>de</u> aldrig* ('<u>they</u> never pay this'), in this cycle (line

15) the information has been reformulated as *den kan jeg ikke betale* (‘I can’t pay this’). Such deictic changes (Steffensen 2012; cf. Bang and Døør 2007) are a cost-efficient way to evoke other parties’ perspectives, in this case that of the invoice receiver. These deictics “refer directly to the personal, temporal or locational characteristics of the situation” (Crystal 2008, p. 133), and, by so doing, anchor verbal/symbolic dimensions (what has been called “the second order language,” cf. Thibault 2011) in real-time speech or first-order languaging. Crucially, then, the result can be traced to neither what is said nor the accompanying non-verbal behaviour: rather it depends on how parties orient to each other with respect to both what is said and how it is that they move. Interactivity, in other words, is sense-saturated in that it links the sense of the deictics (and other words) actually spoken to the pico-scale movements (e.g., shifts in gaze, how syllables are articulated) that are the primordial basis of interactivity.

In short, the DCS uses how deictic markers are integrated with movements and shifts in attention to prompt the participants to reframe the problem. Of course, it would be unwarranted to suppose that the deictics/movements *in themselves* overrule the embodied and situated participant point-of-view. In order to demonstrate the importance of the pico-scale, I examine the 1400 ms of interactivity during which Black utters line 15. This examination, shown in Figs. 5.6 and 5.7, reveals how the participants’ interbodily dynamics are enmeshed with the verbal patterns.

As shown in Fig. 5.7, Black’s utterance act falls into three intonation groups (here rendered in standard orthography):

*hvis det var ˈ*mig | *så røg den ˈbare* | *hen i ˈstakken*
‘if it were me’ | ‘then it just went’ | ‘in the pile’

The most prominent syllables ˈ*mig* ([ˈmai], ‘me’), ˈ*bare* ([ˈbaːɑ], ‘just’), and ˈ*stakken* ([ˈsdɑgŋ], ‘the pile’) function as the Zeitgeber of interbodily dynamics. If we are to understand the dissolution of the fixation bias, it is important to attend to how the parties use their timing. Thus, when Black starts his utterance, he is holding the invoice in front of him in his left hand. As he starts on his first intonational group, Black makes a slight upward movement in his left hand; then, he lets go of the paper, and catches it again, app. 3–4 cm closer to the table. The movement is *perfectly synchronised* with the prosodic pattern: the upward movement co-occurs with the first syllable of the utterance, *hvis* (‘if’), and the catch of the paper is perfectly synchronised with the stressed syllable *mig*, spoken as his thumb touches the paper (130 ms into a stressed syllable that lasts 200 ms) just as his voice begins to fall away.[5] In short, as Black adopts an altero-centric deictic system, his body simulates (cf. Kirsh this vol.) a receiver of the invoice. Quite literally, he receives (catches) the invoice at *exactly the same time* as he utters the deictic *mig* (‘me’).

---

[5] The margin of error is 40 ms, or the time between two frames. The spectrogram (made in PRAAT software) shows with accuracy when the syllable starts, and this can be imposed on the frame-by-frame video annotation (made in ELAN software). This procedure shows that the hand movement *at latest* starts in the frame that immediately follows the frame in which the syllable starts.

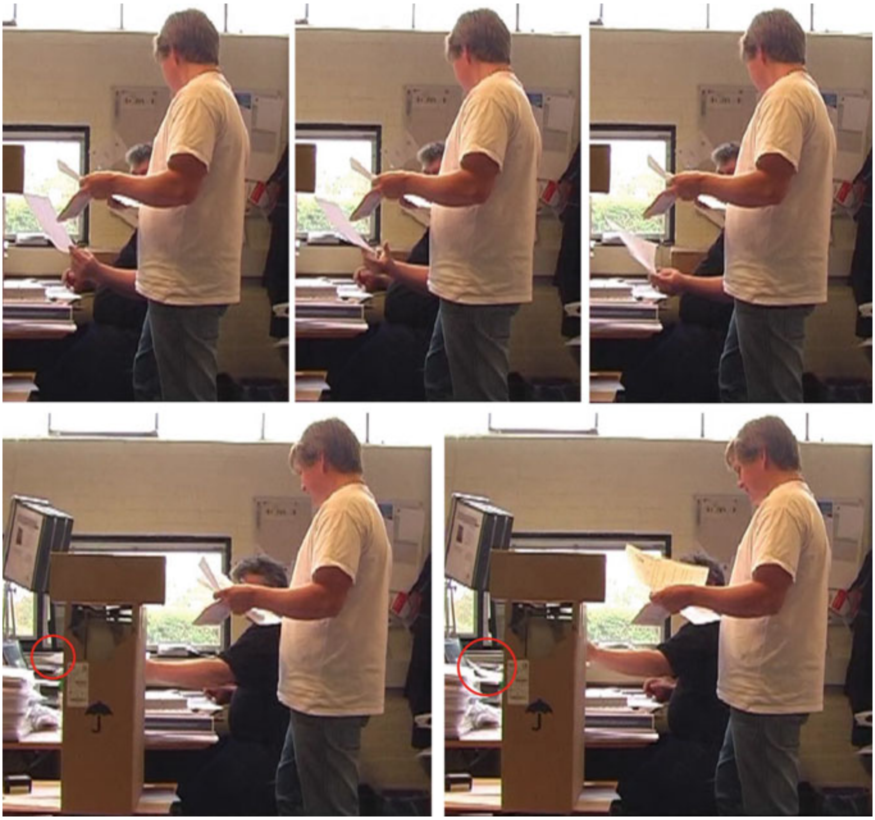


Fig. 5.6 Black embodies the receiver. Picture 1 (−8500): Black elevates the invoice slightly. Picture 2 (−8340): Black lets go of the paper and lowers his hand. Picture 3 (−8140): Black catches the invoice. Picture 4 (−7140): Black leans forward and stretches his arm so the invoice (indicated by the *red circle*) approaches the pile. Picture 5 (−6940): Black lets go of the invoice (indicated by the *red circle*) and lets it drop into the pile

The narrative structure thus prompts him to mimic a receiver in a way that turns his body into a cognitive resource for solving the problem. Black's activity is sense-saturated not only by hypothesising 'if it were me' but also by using deictic recalibration to impose behavioural order.

During the next two intonation groups, Black-as-receiver dissolves the fixation bias based on the print paper version of the invoice. The second intonation group—*så røg den* '*bare* ('then it just went')—pivots on 150 ms long stressed syllable '*bare* ('just'). As he utters this syllable, his bodily dynamics change: having caught the invoice with his left hand, its downward movement becomes a forward-reaching movement that spreads through his body. He then leans forward during the 500 ms of the third intonation group, and as the last syllable ends, lets the paper fall into a

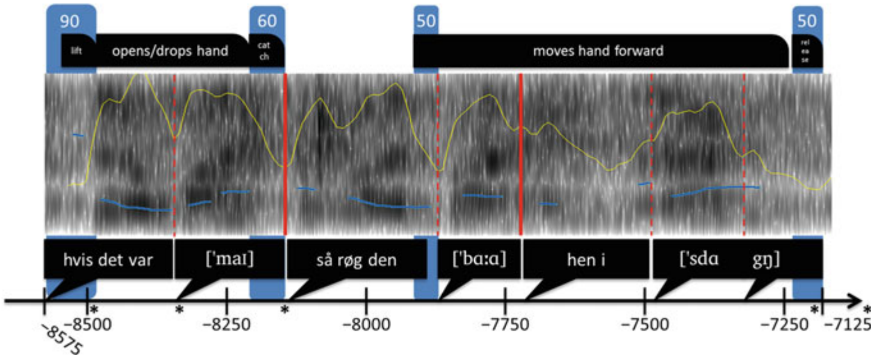


Fig. 5.7 The timeline of line 15 (from −8500 to −7125), corresponding to Fig. 5.6. The *top row* shows Black's hand/arm movements (*squares with a rounded corner*), and the *bottom row* shows Blacks wording with prominent syllables rendered in IPA (in *squared speech bubbles*). Between the two rows, a *spectrogram* shows some of the acoustic properties of Black's utterance act in line 15, incl. pitch (*blue line*) and loudness (*yellow line*). The *blue areas mark* synchronicities in speech/hand coordination: the beginning of the utterance act and the initiation of the invoice handling (90 ms); the prominent syllable [ˈmaɪ] and the catch of the invoice (60 ms); the prominent syllable [ˈbɑːɑ] and the initiation of the forward movement of hand/arm (50 ms); the end of the last syllable [gŋ] and the release of the paper at the end of the forward hand movement (50 ms). The asterisks at the timeline indicate the five points in time that correspond to the five pictures in Fig. 5.6

pile of paper in the far end of the desk. Remarkably, halfway through Black's *mig* ('me'), White turns his head and follows the invoice with his gaze until Black lets go of the paper. As he does so, White's head movement continues and his gaze rests at the papers in his hand. Black's bodily dynamics set off a form of interbodily synchrony. This pattern occurs as Black enacts the movement of the imagined invoice receiver who, having realised that he cannot pay the invoice as it carries no CVR number, throws it into a pile of paper; as he does it, White observes a scenario of what it looks like when the receiver receives the invoice.

Simulating the receiver has direct implications for the DCS. As argued in Steffensen et al. (2010), a simulation entails a doubling in functional levels of the DCS: it both functions as the *simulation* (where Black plays the role as receiver) and as the *situation* (where Black is himself). Obviously, the discard of the print paper invoice plays out as part of the simulation (it is the receiver who throws it away), but the *effects* of the discard are detectable in the situation (it is Black and White who are no longer biased by the print paper invoice). Thus, the whole-bodied achievement of line 15 recalibrates the DCS which prompts it to solve the problem 7100 ms later. This happens as the two participants move out of the narrated receiver's point-of-view (in line 15–17) back to their own. Thus, in line 18 Black responds to his own question without personal deictics (*Der er ikke noget CVR nummer på*, 'There is no CVR number on it') and, in line 19, with a deictic *du* ('you'/'one') that is ambiguous between second person and a generalized 'zero' person (cf. Bang and Døør 2007). In short, it is not clear if he is addressing what

White can and cannot do or some generic rule of invoice-sending. On either interpretation, Black speaks from an ego-centric frame of reference that restores his perspective as the *sender* of the invoice. All in all, this third cycle takes the two participants to an alter-centric frame of reference and back: this cognitive reframing prompts them to dissolve their fixation bias.

Problem-Solving and Solution-Probing

Black has done most of the cognitive labour in the pre-pivot phase. Remarkably, it is White who formulates the solution to the problem in line 20:

20. W: nå nej men det er der jo hvis vi printer ud på logo papir. (0.3)
21. B: er der. (0.7)
22. W: ↑ja ja (1.2) selvfølgelig (0.7)
23. B: nå ja (0.3) det er rigtigt (0.4)
24. W: ja ja (0.6) den er god nok jo.
25. B: ja det er rigtigt
20. W: well no but it [the number] is there if we print on logo paper. (0.3)
21. B: there is? (0.7)
22. W: ↑yes yes (1.2) of course (0.7)
23. B: oh yes (0.2) that's right (0.4)
24. W: yes yes (0.6) it's true, all right (0.4)
25. B: yes that's right

Though Black has done the pre-solution work, White now becomes the *main cognizer* (Galosia et al. 2010) by formulating the solution: he says that the number will appear if the invoice is printed on the logo paper. But *how* does he reach the insight or, perhaps, how does the insight reach him? Though no definite conclusion can be reached, it seems clear that this too depends on sense-saturated coordination (as well as neural events). White's bodily actions suggest that the 'insight' arises through four stages as shown in Fig. 5.8: first, the insight comes 'like a bolt from the blue' at −4500; second, he articulates his insight at 0 (the event pivot); third, he realizes that the problem has been resolved by the proposed solution (at +4200); and fourth, he outlines consequences of adopting the proposed solution (from +5100 onward).

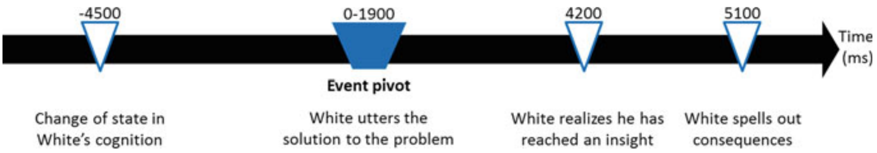


Fig. 5.8 The trajectory of White's insight (from −4500 to +5100). The figure shows the four stages that the DCS undergoes as White reaches the insight

It may seem surprising to claim that White's insight is manifest 4500 ms before he puts it into words. How do we know that he has reached an insight when he has not said so? Close examination of White's posture, head position and gaze, reveals that from −45000 to −4450, he is in constant motion: he moves slightly back and forth, points to the invoice with the pile of papers in his right hand, fiddles with those in his hand, glancing occasionally at them as his gaze follows the invoice in Black's hand. But at −4450 White suddenly looks up from his paper, gazes towards Black's computer screen, and, during Black's utterance act in line 18–19, stands completely motionless for 4,000 ms. Though we cannot observe any cognitive work, we can assume that *neurally* something is happening. Briefly, the DCS pivots on a brain that comes up with the logo paper hypothesis. This change occurs 560 ms after Black ends his alter-centric dissolution of the fixation bias (in line 15–17): this short time frame indicates a direct link between Black's narrative and White's insight.

As White utters the insight in line 20, he turns his head and gazes at Black's computer while pointing at the screen. Strikingly, White gives no sign of noticing that he has solved the problem. The first indication appears when White, at +3230, responds to Black's *er der?* ('there is?'). He now uses a rising intonation on *ja ja* ('yes yes', line 22) that, by +3800, merges into a clear indication that he has grasped the implications of what he has said. White closes the pile of papers, looks up, and leans forward as he stretches his left arm to pick up the invoice from the pile at Black's table. In the midst of the bodily movement, as he utters the prominent syllable [ˈfø] in *selvfølgelig* ([sɛˈføli] 'of course', line 22), he hesitates for 400 motionless ms. Again, it seems that the DCS has come to pivot on White's neural resources as he grasps the implications of a solution.

If it is indeed the case that White utters the solution at 0, and realizes that he had done so at +4200, it implies that the two participants did not exploit language as a tool for aligning states of mind through externalisation, as assumed by mainstream linguists and cognitive scientists. Rather, the wording in line 20 functions as a *hypothesis generator* beyond deductive logic and problem-solving strategies. Thus, rather than engaging in problem-solving, the two participants depend on *solution-probing*: they project probes until they, *post festum*, observe that one of the probes fits the problem space. The participants are thus not merely agents, but also observers of their own interactivity. It is this probing strategy that gives the cognitive trajectory its chaotic, self-organising quality: it does not follow a pre-destined scheme or blueprint; it is messy and meshed with real-time interbodily behaviour.

The post-pivot period is characterized by White's refocusing attention on the invoice which he grasps and places in front of Black. As they have now dissolved the fixation bias, the printed invoice no longer represents their impasse; it is, rather, a resource for completing a workable invoice. In the post-pivot period, they undertake excessive confirmations (line 22–25) that align the two participants in positive emotional resonance that contrasts with the pre-pivot frustration. They resonate not only through verbal affirmations but also in pico-scale interbodily dynamics. This resonance occurs when White picks up the paper and places it on Black's table. As he does so, Black synchronises with White's movement by picking up a pile of papers, and stomping them in the table. Not only is the movement similar: both participants lift their respective papers from a horizontal position to touch the table in a vertical one. What is most striking is that the movements are almost completely synchronous, as shown below in Figs. 5.9 and 5.10.

Black's hand movement starts 325 ms after White's, and strikingly, he grasps his pile of papers 145 ms before White. As Black's hand is closer to his paper than is White's to the invoice, Black makes a circuitous movement: he lifts his hand and points his index finger upwards (in a *eureka*-like way), that is synchronised with a *nå ja* ('oh yes') in line 23 before he proceeds by lowering his hand towards the papers. Within 145 ms both participants lift their respective sheets, move them in

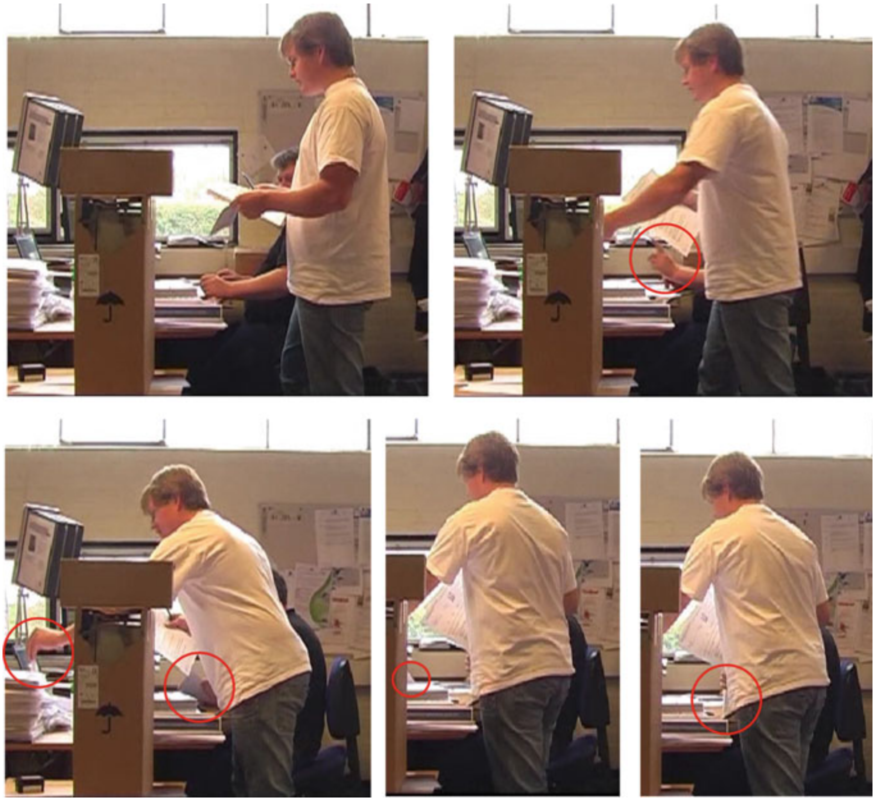


Fig. 5.9 Black and White synchronizes a pick-up of papers. Picture 1 (+5050): Immediately prior to the sequence. Picture 2 (+5650): White has started leaning forward, reaching for the invoice in the pile; Black lifts his index finger (marked by *red circle*). Picture 3 (+7010): White has grasped the invoice and starts lifting it (marked by *red circle*); Black has grasped the papers at the table and starts lifting them (marked by *red circle*). Picture 4 (+8130): White lowers the invoice which is just about to touch the table (marked by *red circle*). Picture 5 (+8290): The papers in Black's hand (marked by *red circle*) have just made contact with the table, 100 ms after the invoice in White's hand did the same

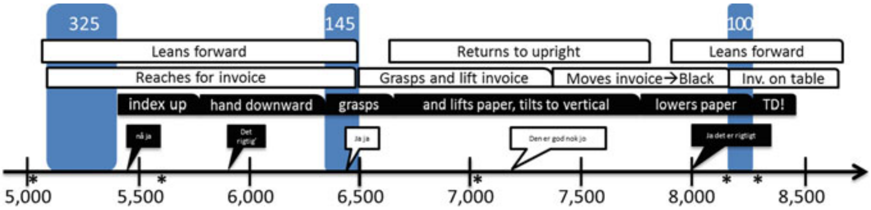


Fig. 5.10 The timeline of the interbodily synchronisation in Fig. 5.9 (from +5000 to +8600). The *four rows* show: White's posture, White's arm/hand, Black's arm/hand, and Black's and White's speech. The synchronisation is marked with *blue*: first when the initiate the hand movements (within 325 ms), second when they pick up the invoice/the papers (within 145 ms), and third when they make the invoice/the papers touch the table (within 100 ms). The *asterisks* at the timeline indicate the five points in time that correspond to the four pictures in Fig. 5.6

an arc, and make the bottom edge of the paper touch the table within 100 ms of each other. On the pico-scale of interbodily dynamics, they engage in a spontaneously choreographed ballet where their papers are moved in a resonant, synchronous pattern that resonates with their (by Danish standards) excessive verbal agreement.

Conclusion: Interactivity, Language and Cognition

In the previous section we have undeniably seen a problem being solved. Loosely speaking, we could say that Black and White solved the problem, more technically that it was solved by a DCS. In terms of the components hierarchy suggested above, the DCS under scrutiny included human beings, artefacts, and narratives. Obviously, a large part of the cognitive labour consisted in how the interbodily dynamics of the two participants iteratively calibrated and aligned the DCS. These dynamics comprised: (i) Sharing of emotional states (e.g., frustration before the event pivot and exultation after); (ii) timing of vocal gestures (e.g., syllabic prominence as a Zeitgeber for bodily actions); (iii) coordination of gaze: partly as shared attention at the invoice, partly as eye contact; (iv) synchronisation of movement (e.g., in how papers are handled).

Especially the handling of the print paper invoice played a vital part in the cognitive event trajectory. Thus, the invoice became a main artefact for Black as he reframed the problem through a simulation of how the receiver would handle the invoice. Significantly, the invoice undertook a spatial reorganisation, as the two participants actively turned their environment into a cognitive resource. As indicated in Fig. 5.11, the spatial reorganisation correlated with the event trajectory: the invoice both starts (until −18625) and ends (from +6500) on the table in front of

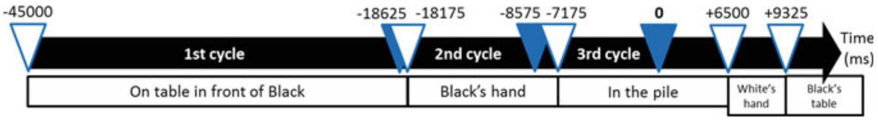


Fig. 5.11 The trajectory of the print paper invoice throughout the cognitive event. The figure shows the spatial reorganisation from the table to the pile and back, mediated by Black's and White's handling it, respectively

Black, first as a fixation bias, then as a resource in their future work. Between these two points, the invoice is placed in the pile at the far end of the table. Strikingly, Black puts it there, while White takes it back.

The active, real-time, spatial manipulation of artefacts thus contributes to the cognitive event trajectory. One could argue that the spatiality of the DCS environment is thus a main component that should be added to the hierarchy established in section "Components in Distributed Cognitive System". This would be in line with Hollan et al.'s (2000:179) observation: "To understand human cognition, it is not enough to know how the mind processes information. It is also necessary to know how the information to be processed is arranged in the material and social world." The importance of the spatial dimensions also came to the fore when Black scanned his noticeboard as part of the environment in search for the CVR number.

The narrative contributed to the DCS through its ability to prompt the participants to see the problem from other points of view, *in casu* the view of the receiver of the invoice. This was achieved partly through the narrative formulaic *hvis det var mig* ('if it were me'), partly through the use of personal deictics with an alter-centric reference. All in all, the wording and the narrative prompted the participants to reframe the cognitive problem at hand.

The analysis also gave rise to another insight into the nature of problem-solving. Far from being a testimony of human rationality, the analysis showed a modus operandi that does not depend on developed schemes, plans, and blueprints for solving problems. The DCS under scrutiny moved through the problem-space along a cognitive trajectory that was self-organised, unpredictable and on the edge of chaos. We did not witness analytical problem-solving, but rather creative *solution-probing*. The probing nature of how Black and White solved the problem points in two directions. First, it seems that the pervasiveness of social institutions that operate in "highly rationalized environments" (McGarry 2005, p. 187)—including science, at least as it presents itself front-stage—has biased our understanding of how human beings solve problems. Such institutions depend on standardised and automatised modes of working where norms and procedures impose an orderliness that turns creative solution-probing into analytical problem-solving. However, even in such rationalised environments, the DCS has an underlying capacity for self-organising cognitive processes in a way that phenomenologically appears to us as *creativity*, *intuition*, and *Aha!* Second, it seems that standard versions of theories on human problem-solving depend on analytical and cognitive psychological models of processing input (the problem) in order to generate output (the solution). This calls for a theoretical reassessment of the embodied and interbodily dynamics of human problem-solving, a reassessment that upholds the procedural model of problem-solving as a subset of a more general human capacity. Such a subset comes in handy in complex work environments where speed or security become decisive, for example in emergency medicine (cf. Pedersen 2012) or in aviation (cf. Desai et al. 2012), but even in such environments intuition and creativity can be crucial elements in achieving a successful result.

This chapter has made two contributions to a more realistic view on human problem-solving. First, it devised a method of *Cognitive Event Analysis*. Most saliently, this method pivots on scrupulously close examination of video-recorded events (building on such work as Cowley 1998; Steffensen et al. 2010; Thibault 2011; Pedersen 2012). The method thus allows for detailed scrutiny of *what happens* as cognitive results are brought forth. It consists of two steps, an event identification procedure, and an event trajectory analysis (as mentioned, it has later been expanded to comprise five steps, cf. Steffensen et al. 2016). It thus combines a deductive procedure with an inductive approach to data. While the former allows the method to integrate in various trans-disciplinary fields (e.g., the study of cognitive events in the workplace, in aviation, or in health settings), the latter appreciates Clark's (2008, p. 13) *Principle of Ecological Assembly*. Thus, if the enabling conditions of problem-solving include "whatever mix of problem-solving resources [that] will yield an acceptable result," there is no a priori way of delineating what ought be to investigated as contributing to the emergence of a solution. As argued, the whole array of situated elements, including interbodily dynamics and artefacts, may contribute to the cognitive event, but the same goes for many non-situated elements, including sociocultural resources, verbal patterns, narratives, memorised procedures, and autobiographical memory. Such non-situated components of a DCS pose a methodological problem for observational methods, because they depend on the non-situated parts being situated, and that does not always happen in an overt way. However, if the observational data are complemented by careful ethnographic methods—for example, interviewing and participant observation—it is to an extent possible to counter this shortcoming. Likewise, a careful inclusion of experimental data may prove fruitful: Thus, if a given structure is known to yield cognitive results in controlled experimental settings, the same structure is, *ceteris paribus*, likely to yield similar results in the wild.

The second contribution of the current chapter consisted in an interactivity-based approach to problem-solving. The present analysis points to the necessity of grounding the study of cognition in interactivity, or sense-saturated coordination that contributes to human action. In allowing interactivity to play the decisive role in human problem-solving, we remove some of the cognitive burden from the brain.

This view on cognition contrasts with what we can term the “dermal metaphysics” of traditional cognitive science, sociology, and biology, i.e. the view that the skin is the absolute barrier between two distinct realms: the internal versus the external, biology versus sociality, cognition versus communication, or meaning versus behaviour. With interactivity, we need not make any such a priori distinctions: They all come together in what people do as they achieve results. Thus, interactivity is an ontological substrate that can be described both as cognition, as language, as ecological niche-construction, and as behaviour. Accordingly, we must discard “(T)he simplemindedness of mainstream cognitive psychology about the nature and function of human communication” (Reed 1996, p. 157): language is *not* an instrument for externalising thoughts, language is *not* an instrument for exchanging pre-existing meaning, and language is *not* a pre-existing system that we can “use.” Rather, languaging is meshed with real-time interbodily behaviour. It functions as a Zeitgeber for action (cf. Steffensen et al. 2010) and as a constraint on pico-scale behaviour. The latter is exemplified by Black and White when they exploited the symbolic dimensions of languaging to simulate the receiver’s point of view. In this way, languaging enables participants to perform low-cost categorisation and simulation in their ecology (cf. Clark 2008, Chap. 3). The mechanism that allows us to do this is our capability to take a *language stance* (Cowley 2011). The language stance allows us to control our pico-scale bodily behaviour (e.g., as we produce and perceive specific vocal gestures) in a way that conforms to our phenomenological experience of engaging with verbal patterns.

Finally, it should be pointed out that an interesting implication of an interactivity-based approach is that it links to the philosophical program of Ludwig Wittgenstein (1953). Just as Wittgenstein sought to move beyond, or below, linguistic meaning with recourse to what people do with language, the interactivity approach seeks to move below the phenomenology of language by exploring human behaviour on a pico-scale. On the micro-scale of wording and social interaction, people can agree (or not): they can align verbally, phenomenologically and consciously. On the pico-scale, people can resonate and synchronize (or not)—they can align sub-consciously, sub-phenomenologically—or just interbodily. With Wittgenstein, we can say that on the pico-scale there “is not agreement in opinions but in form of life” (Wittgenstein 1953: 88 [§241]).

Acknowledgements I am grateful to the editors for the valuable comments and suggestions. Likewise, Sarah Bro Trasmundi (University of Southern Denmark) contributed with helpful suggestions, and Rasmus Nielsen (University of Southern Denmark) assisted in the interpretation of the spectrogram in Fig. 5.7. Finally, I am indebted to Anne-Kathrine Brorsen who allowed me to use her data for the in-depth analysis, and not least indebted to Black and White for sharing their problem-solving with us.

References

Ball, L. J., & Litchfield, D. (this vol.). *Interactivity and embodied cues in problem solving, learning and insight: Further contributions to a "theory of hints"*.

Bang, J. C., & Døør, J. (2007). *Language, ecology and society: A dialectical approach*. In S.V. Steffensen & J. Nash (Eds.). London: Continuum.

Berkowitz, L. (2007). Frustration-aggression hypothesis. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 366–368). Thousand Oaks, CA: Sage.

Brorsen, A.-K. (2010). *Til dialog…* Unpublished elite module project. Odense: University of Southern Denmark.

Bråten, S. (2009). *The intersubjective mirror in infant learning and evolution of speech*. Amsterdam: John Benjamins.

Chemero, A. (2000). What events are: Commentary on Stoffregen's "Affordances and events". *Ecological Psychology, 12,* 37–42.

Clark, A. (2008). *Supersizing the mind: Embodiment, action and cognitive extension*. Oxford: Oxford University Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58,* 7–19.

Cowley, S. J. (1998). Of timing, turn-taking, and conversation. *Journal of Psycholinguistic Research, 27,* 541–571.

Cowley, S. J. (2011). Distributed language. In S. J. Cowley (Ed.), *Disitributed language* (pp. 1–14). Amsterdam: John Benjamins.

Cowley, S. J. (2012). Mimesis and language: A distributed view. *Interaction Studies, 13,* 17–40.

Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th ed.). Oxford: Blackwell.

Desai, V., Madsen, P. M., & Roberts, K. H. (2012). High-reliability organizations in health care. In P. Carayon (Ed.), *Handbook of human factors and ergonomics in health care and patient safety*. Boca Raton FL: CRC Press, Taylor and Francis Group.

Dollard, D., Doob, L., Miller, N., Mowrer, O., & Sears, R. (1939). *Frustration and aggression*. New Haven, CT: Yale University Press.

Donald, M. (1991). *The origins of the modern mind*. Cambridge MA: Harvard University Press.

Donald, M. (2001). *A mind so rare*. New York: W.W. Norton.

Donald, M. (2012). Commentary on Cowley's chapter. *Interaction Studies, 13,* 41–49.

Galosia, M., Cowley, S. J., Aliner, G., & Steffensen, S. V. (2010). When life hangs on biosemiosis: Tracking expertise in a medical simulation centre. In S. J. Cowley, J. C. Major, S. V. Steffensen, & A. Dinis (Eds.), *Signifying bodies: Biosemiosis, interaction and health* (pp. 173–200). Braga: Portuguese Catholic University Press.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32,* 1489–1522.

Harel, I. (1991). The silent observer and holistic note taker: Using video to document a research project. In I. Harel & S. Papert (Eds.), *Constructionism*. Norwood, NJ: Ablex Publishers.

Hemmingsen, B. (2013). The quick and the dead: On temporality and human agency. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Interactivity, computation and human artifice*. Dordrecht: Springer.

Hodges, B. H., & Baron, R. M. (1992). Values as constraints on affordances: Perceiving and acting properly. *Journal for the Theory of Social Behaviour, 22,* 263–294.

Hollan, J., Hutchins, E., & Kirsch, D. (2000). Distributed cognition: Toward a new foundation for human–computer interaction. *ACM Transactions on Human-Computer Interaction, 7,* 174–196.

Hutchins, E. (1995a). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science, 19,* 265–288.

Järvilehto, T. (1998). The theory of the organism-environment system: I. Description of the theory. *Integrative Physiological and Behavioral Science, 33,* 321–334.

Järvilehto, T. (2009). The theory of the organism-environment system as a basis of experimental work in psychology. *Ecological Psychology, 21,* 112–120.

Kirsh, D. (1997). Interactivity and multimedia interfaces. *Instructional Sciences, 25,* 79–96.

Kirsh, D. (2006). Distributed cognition: A methodological note. *Pragmatics & Cognition, 14,* 249–262.

Kirsh, D. (this vol.). *Thinking with external representations*.

Linell, P. (2009). *Rethinking language, mind, and world dialogically: Interactional and contextual theories of human sense-making*. Charlotte, NC: Information Age Publishing.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Sciences, 20,* 130–141.

Maier, N. R. F. (1930). *Problem solving and creativity in individuals and groups*. Belmont, CA: Brooks.

McGarry, B. (2005). *Things to think with: Understanding interactions with artefacts in engineering design*. Ph.D. Dissertation, University of Queensland, St Lucia.

Pedersen, S. B. (2010). Towards dialogical health care practices: Human errors as a result of cultural affordances. In S. J. Cowley, J. C. Major, S, V. Steffensen, & A. Dinis (Eds.), *Signifying bodies: Biosemiosis, interaction and health* (pp. 245–276). Braga: Catholic University of Portugal.

Pedersen, S. B. (2012). Interactivity in health care: Bodies, values and dynamics. *Language Sciences, 34,* 532–542.

Pedersen, S. B. (2015). *The cognitive ecology of human errors in emergency medicine: An interactivity-based approach* (Ph.D.). Odense: University of Southern Denmark.

Perry, M. (this vol.). *Socially distributed cognition in loosely coupled systems*.

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. Oxford: Oxford University Press.

Robbins, P., & Aydede, M. (Eds.). (2009). *The Cambridge handbook of situated cognition*. Cambridge: Cambridge University Press.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50,* 696–735.

de Saussure, F. (1916/1972). *Cours de linguistique générale*. Paris: Payot.

Scheerer, M. (1963). Problem solving. *Scientific American, 208,* 118–128.

Steffensen, S. V. (2009). Language, languaging and the extended mind hypothesis. *Pragmatics & Cognition, 17,* 677–697.

Steffensen, S. V. (2011). Beyond mind: An extended ecology of languaging. In S. J. Cowley (Ed.), *Distributed language* (pp. 185–210). Amsterdam: John Benjamins.

Steffensen, S. V. (2012). Care and conversing in dialogical systems. *Language Sciences, 34,* 513–531.

Steffensen, S. V. (2015). Distributed language and dialogism: Notes on non-locality, sense-making and interactivity. *Language Sciences*, *50*, 105–119. doi:10.1016/j.langsci.2015.01.004.

Steffensen, S. V., & Cowley, S. J. (2010). Signifying bodies and health: A non-local aftermath. In S. J. Cowley, J. C. Major, S. V. Steffensen, & A. Dinis (Eds.), *Signifying bodies: Biosemiosis, interaction and health* (pp. 331–356). Braga: Catholic University of Portugal.

Steffensen, S. V., Thibault, P. J., & Cowley, S. J. (2010). Living in the social meshwork: The case of health interaction. In S. J. Cowley, J. C. Major, S. V. Steffensen, & A. Dinis (Eds.), *Signifying bodies: Biosemiosis, interaction and health* (pp. 207–244). Braga: Catholic University of Portugal.

Steffensen, S. V., Vallee-Tourangeau, F., & Vallee-Tourangeau, G. (2016). Cognitive events in a problem-solving task: A qualitative method for investigating interactivity in the 17 animals problem. *Journal of Cognitive Psychology*, *28*(1), 79–105. doi:10.1080/20445911.2015.1095193

Thibault, P. J. (2011). First order languaging dynamics and second-order language: The distributed language view. *Ecological Psychology, 23,* 210–245.

Trevarthen, C. (1998). The concept and foundations of intersubjectivity. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 15–46). Cambridge: Cambridge University Press.

Weisberg, R. W., & Alba, J. W. (1981). An examination of the alleged role of “fixation” in the solution of several “insight” problems. *Journal of Experimental Psychology: General, 110,* 169–192.

Wertsch, J. (2002). *Voices of collective remembering*. Cambridge: Cambridge University Press.

Wilson, R. A., & Clark, A. (2009). How to situate cognition. Letting nature take its course. In Robbins, P., & Aydede, M. (Eds.), *The Cambridge handbook of situated cognition.* Cambridge: Cambridge University Press.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Basil Blackwell.

Chapter 6
Interactivity and Embodied Cues in Problem Solving, Learning and Insight: Further Contributions to a "Theory of Hints"

Linden J. Ball and Damien Litchfield

Abstract This chapter addresses the situated, embodied and interactive characteristics of problem solving by focusing on the cues that arise within a solver's external environment. In examining the influence of external cues on problem solving we have been heavily influenced by Kirsh's (The Cambridge handbook of situated cognition, Cambridge University Press, Cambridge, 2009) "theory of hints". We extend this theory to include hints that derive from the communicative properties of other people's eye movements, focusing on the role of eye gaze in directing attention and conveying information that can be beneficial for problem solving. A particularly interesting aspect of eye gaze is its capacity to facilitate the perceptual priming of motor simulations in an observer. This gives rise to the potential for an expert problem solver's eye movements to cue imitative perceptual and attentional processing in less expert observers that can promote effective problem solving. We review studies that support the hypothesised role of gaze cues in scaffolding problem solving, focusing on examples from insight tasks and diagnostic radiography. Findings reveal that eye gaze can support a variety of decisions and judgments in problem solving contexts. In sum, knowing where another person looks provides hints that can act both implicitly and explicitly to cue attention and to shape thoughts and decisions.

There are many occasions when we find that we have a problem to solve, that is, we have a goal that we want to achieve but do not immediately know what to do to reach that goal. Such problems can be fairly mundane, such as trying to keep dry when you get caught in an unexpected downpour or trying to find your way around a new city that you are visiting for the first time. The problem continuum can also stretch to more profound goals, such as working out what you need to do to avoid

---

L.J. Ball (✉)
University of Central Lancashire, Preston, UK
e-mail: LBall@uclan.ac.uk

D. Litchfield
Edge Hill University, Ormskirk, UK

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_6

bankruptcy or striving to resolve a major scientific research question. An issue that is intimately connected to problem solving is that of "transfer", which is primarily concerned with the benefits that prior experience and knowledge can bring to current problem solving. The transfer theme connects very closely with topics such as learning and the development of expertise, since there is much research showing that as individuals engage deliberatively in tackling problems within a particular domain they transition from novice to expert status through processes of knowledge acquisition and knowledge restructuring (e.g., Van De Weil et al. 2000) as well as through the development of highly effective storage and retrieval mechanisms to support the use of domain-specific knowledge (e.g., Ericsson and Kintsch 1995). Indeed, when advanced levels of expertise are attained many tasks within a domain may end up becoming fairly routine, no longer having the status of traditional problems, since the experienced practitioner not only possesses the requisite knowledge to solve the task but also possesses the methods that are needed to ensure the effective retrieval and application of such knowledge.

In this chapter we begin by considering some of the changing theoretical conceptions of problem solving and learning that have arisen from a growing appreciation that these activities need to be understood as being fully situated, embodied and interactive in nature (e.g., Kirsh 2009), as opposed to the classical view (e.g., Newell and Simon 1972), whereby problem solving is decontextualised, disembodied and divorced from external resources that are present within the solver's physical and cultural environment. The present chapter will touch upon a variety of important issues that relate to the situated, embodied and interactive characteristics of problem solving as conceptualised in contemporary research, with a key focus throughout our discussion being placed on the important role played by cues that arise within the solver's external environment.

In this latter respect we agree fully with Kirsh's (2009) assessment that classical problem solving theory has failed adequately to grapple with "the universality of cultural products that facilitate activity-specific reasoning" (Kirsh 2009, p. 284). What Kirsh is referring to here is the idea that the environments in which people regularly act are replete with "mental aids", such that problem solving becomes more a matter of making effective use of these available aids than relying purely on existing knowledge and internal cognition. As Kirsh goes on to say, "Our problems arise in socially organized activities in which our decisions and activity are supported" (p. 284). Such support structures take the form of "scaffolds" and "resources" that are designed to make it easier for people to complete their tasks, whether these are wayfinding problems, problems in using technological devices and appliances or problems in choosing goods in a local supermarket to meet nutritional preferences and budgetary constraints. Scaffolds and resources can, more generally, be viewed as "hints", with a key source of such hints being our colleagues, neighbours, supervisors, teachers, trainers and the like, who are usually at hand to offer assistance by providing helpful suggestions, clues, advice and tools. Kirsh (2009) views such hints as providing a key basis for an alternative and positive theory of how people overcome problems in concrete settings through a dynamic process of agent-environment interaction.

We have taken inspiration from Kirsh's (2009) recent sketch of a "theory of hints", and in the present chapter we attempt to provide a small addition to this outline theory, albeit an addition that we believe is a vitally important one. Our particular focus is on the role of hints that derive from the communicative properties of others' eye movements in directing attention and conveying information (e.g., Kleinke 1986). The question underpinning our research is a simple one: if a person's eye movements can direct our attention, then can we also learn to make use of these cues wherever possible in a way than can shape and facilitate our problem solving? This is an important empirical question that we and other contemporary researchers are actively addressing, and one of the major goals of this chapter is to review some of the key literature in this area. By way of pre-empting our answer to the question, we note up-front that the existing evidence points resoundingly to an affirmative conclusion. We further suggest that such a positive answer also brings with it the need to augment a theory of hints in a way that accommodates the problem solving scaffolds than come from our inherent human sensitivity to other people's eye gaze.

Changing Conceptions of Problem Solving and Learning

The core focus of much traditional research on problem solving, learning and expertise up to the 1990s was on relatively small-scale tasks being undertaken by individuals working within highly controlled laboratory conditions. Well-defined "puzzle" problems (e.g., the Tower of Hanoi task, river-crossing problems) and games (e.g., chess, tic-tac-toe) featured heavily in this research endeavor and formed the "fruit-flies" of much influential early theorising, inspiring the classical theory of problem solving espoused by Newell and Simon (e.g., 1972; see also Simon 1981). This classical theory views problem solving as involving a heuristically-guided search through a representation of the task (i.e., a "problem space") from an initial state to a goal state via intermediate states. This classical view has many strengths, not least the elegance of portraying a highly generalised view of human problem solving as an adaptive process that is finely tuned to a set of environmental constraints—the so called "task environment" of a problem—that can be conceptualised by theorists as forming the core, abstract, structural aspects of the problem. An individual's problem space may represent more or less of this task environment, since there may be omissions or commissions, the latter deriving from prior biasing assumptions that the individual brings to bear. In addition, the problem space may not only comprise a mental representation but may also be distributed over external resources (e.g., Larkin and Simon 1987; Tabachneck-Schijf et al. 1997).

Even during the heyday of this classical approach to explaining problem solving there were many dissenting voices, which came from those who were concerned that studying decontextualised problem solving by lone individuals was missing much of the richness of real-world problem solving. Those clamoring for a re-focusing of the research agenda tended to base their arguments on two major issues. First, the nature of most real-world, problem solving activity means that it is highly "situated", such that cognitive processes are likely to be shaped heavily by organisational and cultural goals, social structures and interpersonal interactions, as well as by the external artefacts, tools and representational systems at people's disposal (e.g., Suchman 1987). As an example, take the problem solving that arises in the domain of professional design practice. Commercial design is typically observed to be heavily bounded by particular company contexts, which means that design processes are influenced by the constraints and affordances that derive from team members and managers, organisational priorities and goals and established cultural conventions (e.g., see Ball and Ormerod 2000a, b; Reid et al. 2000).

Second, real-world problem solving tends to be highly "distributed", in that cognition it is not located within any one individual, but is instead mediated through complex interactivity between multiple internal and external knowledge repositories, including other team members and various external artefacts (e.g., Busby 2001; Lave 1988). One upshot of this "distributed cognition" approach (e.g., Hutchins 1995) is that it brings into sharp focus the inherent poverty of the classical view, in which context and culture merely moderate the internal cognitive processes of individuals. Instead, it is proposed that "cultural activity systems" (Hutchins 1995) can have cognitive properties of their own that reflect emergent aspects of the whole, rather than simply being the sum of the properties of the individuals working within the system.

Somewhat paradoxically, much of the success of the classical theory of problem solving arguably derived by virtue of the way in which context and culture were effectively *controlled out* of laboratory-based experiments, in essence giving a sterilised view of the reality of real-world problem solving As a consequence, what subsequent problem solving research has revealed very pointedly is that there can be major discrepancies between phenomena that arise "in vivo" and phenomena that have been established "in vitro". For example, Dunbar and Blanchette (2001) demonstrate that the wealth of spontaneous analogical reasoning and communication that arises in real-world problem solving all but disappears in laboratory studies, unless people are explicitly instructed to analogise. In a similar vein, Hutchins (1995) observed how the well-established laboratory-based phenomenon of "confirmation bias" (i.e., a tendency for *individuals* to engage in verifying rather than falsifying tests of favoured hypotheses) was absent in team-based hypothesis-testing in a navigational decision-making context. Like Hutchins, we have also provided evidence for another well-known cognitive tendency, "satisficing"—where problem solvers fixate upon a satisfactory solution rather than exploring options to go beyond mere satisfactory outcomes—as being a dominant force in individual design activity, whilst being largely eradicated in the interactivity arising in team-based design practice (Ball and Ormerod 2000b).

Nowadays, few researchers remain wedded to the classical theory of problem solving. Instead, much research attention is now focused on developing theories that fully embrace the situated, embodied and interactive nature of problem solving using methodologies such as cognitive ethnography that are appropriate for capturing and analysing the richness of real-world cognition (see Ormerod and Ball in press). As we noted above, our particular aim in the present chapter is to focus on the interactivity that arises in real-world, distributed problem solving contexts in order to contribute further conceptual insights to Kirsh's (2009) recent sketch of a "theory of hints". Hints take the form of a wide variety of scaffolds and resources that often originate from the people around us (e.g., colleagues, managers, advisors and the like), who provide clues, suggestions and tools. As Kokinov et al. (1997) note, hints provide a crucially important element of the "culture" of problem solving. Given their apparent importance, we agree with Kirsh's (2009) proposal that any theory of situated problem solving also needs to explain why hints are so successful. Likewise, such a theory needs to be able to accommodate the many different ways in which our problem solving environments offer up hints that enable us to tackle problems more effectively.

In outlining what a relatively simple theory of hints might look like, Kirsh (2009) suggests that it should begin by first defining what a hint is, which he takes to be a verbal or nonverbal cue that acts like a heuristic that can bias the search process. This definition is appealing, not least because it aligns in a constructive way with key concepts from classical problem solving theory, whereby problem solving is viewed as heuristically-guided search though a problem space. Of course, contemporary situated accounts of problem solving downplay the notion of problem-space search in favour of concerns with the external environment and the socio-cultural context as well as with the way in which such factors ground cognition. But as Kirsh emphasises, no comprehensive account of problem solving can overlook the vital importance of understanding the way in which candidate solution ideas are "generated" and subsequently "evaluated" by the problem solver. In this way, an analysis of the manner in which hints provide candidate solutions whose adequacy can be tested gives a foundation for something of a rapprochement between the classical and situated accounts of problem solving. Indeed, the upshot of such a reconciliation of views is that the resulting theory has the potential to afford a highly positive account of how people overcome problems in concrete settings through a dynamic process of agent-environment interaction.

Kirsh's theory views hints as having a particularly valuable *generative* function in games such as chess, where there are a vast number of choice points that need to be considered to make an effective move. Thus, typical verbal hints for opening a game include advice such as "Open with a centre pawn" or "Knights before bishops", which serve to bias candidate generation to prudent options. In standard, well-defined "puzzle" tasks, where only a relatively small number of discrete moves are possible, Kirsh suggests that hints are more likely to have an *evaluative* function, helping the problem solver to determine whether a move or action is a good one that has the potential to lead to the desired goal state. For other types of problems, such as those where it is not easy to determine what the available options are or even whether one is at a choice point, Kirsh suggests that hints are most likely to be beneficial if they can help one to *frame* the problem in a constructive manner. In this way, hints may help the problem solver to break away from problem frames arising from the inappropriate application of prior assumptions and beliefs. Such false assumptions may result in the problem solver succumbing to mental "set" or "fixation", which, in turn, can induce a phase of "impasse", where the individual is stuck and cannot make further progress. Indeed, problems that are very hard to solve tend to be of this type, where people find it difficult to escape the shackles of familiar ways of proceeding that are, in fact, inappropriate for solving the current task (Dominowski and Dallob 1995). Such problems are referred to as "insight" tasks (e.g., Gilhooly and Murphy 2005; Kaplan and Simon 1990; Ohlsson 1992), since a successful solution typically only arises as a consequence of a radical reframing or restructuring of the problem that engenders a sudden insight as to the necessary path to a solution (Smith and Kounios 1996).

Hints in Insight Problem Solving

In the case of solving, pioneering studies by Gestalt psychologists have shown how environmental hints that direct attention to a particular item of information associated with the problem setting can assist the problem solver in finding an appropriate problem frame that can lead to an insightful solution. A classic case of the benefits of such environmental hints can be seen in Maier's (1931) study of the two-string problem. In his study, Maier brought the participant into a room containing various objects (e.g., poles, pliers) as well as two strings that were hanging from the ceiling. The participant was then asked to tie together the two strings, but in attempting to do this discovered that the strings were too far apart, such that it was impossible to reach one string while holding onto the other one. The "insight" solution—provided by only a few participants—was to tie the pliers to one of the strings and set it swinging like a pendulum so as to be able to catch it on the upswing while holding the other string. Crucially, Maier found that this insight could be engendered spontaneously by having the experimenter accidentally brush against one of the strings so as to set it swinging, thereby providing a subtle hint that the participant could exploit. Maier argued that participants were not consciously aware of being influenced by the experimenter's action.

Despite the compelling nature of Maier's account, we nevertheless note that the interpretation of his findings is rendered inconclusive because of a failure to use a control condition in which participants received no hint. Without such a control condition it is unclear whether participants benefited simply from having increased exposure to the problem (see Landrum 1990, for a failure to find a hint effect when using a no-hint control group as a comparison condition). Other research, however, has provided convincing evidence that explicit hints operating at a conscious level can facilitate insight into the two-string problem. For example, Battersby et al. (1953) found that reduced solution latencies on the two-string problem could be affected by simply highlighting objects within the room that might be of relevance to a solution.

Results from a recent study by Thomas and Lleras (2009b) have provided some of the most compelling evidence for the role of *non-conscious* hints in guiding insight in the two-string problem. In their study, Thomas and Lleras asked participants to attempt the problem while occasionally taking exercise breaks during which they moved their arms either in a manner related to the problem's solution (the "swing" group) or in a manner inconsistent with the solution (the "stretch" group). Although a majority of participants were unaware of the relationship between their arm-movement exercises and the problem solving task, individuals in the swing group were significantly more likely to solve the problem than individuals in the stretch group.

Thomas and Lleras claim that their findings are consistent with theories of "embodied cognition" (e.g., Barsalou 1999; Gibbs 2006; Spivey 2007; Wilson 2002; Zwaan 1999), which propose that to understand how the mind accomplishes its goals (e.g., solving problems) one has to understand the mind in the context of its links to a *body* that interacts with the physical world. In this way, embodied cognition theories propose that knowledge representations in the brain maintain properties of the sensorimotor states that gave rise to them in the first place (e.g., Barsalou 1999). More recently, Barsalou (2008) has defended a more general account of "grounded cognition", which proposes that much cognition is underpinned by modal simulations (including mental imagery), bodily states and situated actions. Hutchins (2010) evokes the related concept of "enactment" to argue how enacted multimodal representations are involved in the construction of memories for the past, the experience of the present, and the attainment of future goals, as in problem solving.

Thomas and Lleras' (2009b) results with the two-string problem fall neatly within these embodied, enacted and grounded views of cognition, since their findings show how bodily actions can influence thinking, such that people can be guided implicitly toward insightful solutions by directing their actions in solution-appropriate ways. Over the past decade there have, in fact, been a wealth of "demonstration experiments" (Barsalou 2010) within the general area of grounded cognition, which show how grounded mechanism seem to lie at the heart of goal-directed cognition that is related to perception and action, memory, conceptual processing, language comprehension, social judgement and thought (see Barsalou 2010, for a highly focused review of this compelling evidence).

The embodied cognition effects in insight problem solving that have been identified by Thomas and Lleras' (2009b) clearly take the concept of "hints" in an important new direction, since what is being evidenced appears to be the implicit perceptual priming of motor representations in the brain that are associated with "swinging" movements. We still view such embodied priming effects as arising from a form of *hint*, but a hint that is nevertheless a long way removed from an a direct verbal instruction, an explicit gestural prompt or an implicit environmental cue. Our own particular interest is in yet another type of external hint, which is of the kind that can derive from other people's eye movements, where such eye movements can serve to convey information and direct attention and may also provide a basis for imitative motor simulations.

Hints arising from following other people's eye movements are clearly distinct from normal language-based or gesture-based directives, being both more subtle and highly multifaceted in nature. We review the way in which eye movements can act as hints in the next section, where we examine the communicative properties of other people's eye movements, such as their capacity to convey information and to direct our attention (e.g., Kleinke 1986; Sheperd 2010), and where we also entertain the role of eye movements in priming task-relevant perceptual and motor simulations in an observer. Suffice it to say for now that eye movements seem to engender a wide array of cueing effects, ranging from implicit and embodied priming of motor sequences and perceptual simulations right through to explicit and directive communicative prompts.

Hints Arising from Other People's Eye Movements

Many studies have demonstrated that we are highly sensitive to other people's gaze (Gibson and Pick 1963; Symons et al. 2004) and that this sensitivity develops very early in life, with even young infants showing sophisticated gaze following behaviour (Brooks and Meltzoff 2005; Corkum and Moore 1995). Following someone else's gaze can guide attention towards a particular object or item, which in turn enables knowledge to be inferred or to be directly shared, as is the case when observers verbalise their mutual interest in a jointly attended item (Flom and Pick 2007; Hanna and Brennan 2007). Such "joint attention" has been claimed to form a basis for much early learning that arises in development (Butterworth and Jarret 1991), although it also remains ubiquitous in adulthood (Driver et al. 1999). Adult observers are particularly adroit at taking into account nearby objects when following another's gaze, with neuropsychological evidence indicating that following a person's gaze can function to transfer the *intentionality* of that person to the observer (Becchio et al. 2008; Lobmaier et al. 2006). Therefore, by watching where another person looks we alter our own processing of objects. At a very minimum, we obtain cues from where the other person is attending that have the potential to be highly beneficial for our own cognitive activities—such as problem solving—especially if we are in the presence of an individual who has greater domain expertise than we do.

Given the ubiquity of attention-directing eye movements that arise constantly in our everyday work and social environments it would seem that a great deal of ongoing high-level cognition in areas such as problem solving, reasoning, judgment and decision making might be shaped by such external cues. Several recent studies have examined this issue by investigating the value of the direct presentation of the eye movement patterns of one observer to another observer. The aim of these studies is to assess how this form of attentional guidance can improve Performance in a range of visual search and problem solving tasks. Experts are known to look at task-relevant areas more often than novices and also demonstrate more effective search strategies (e.g., Chapman and Underwood 1998; Charness et al. 2001; Krupinski 1996), which means that there is a growing belief that an expert's eye movement patterns should be particularly useful in training novices where to look (Gegenfurtner et al. 2011; Jarodzka et al 2012; Krupinski et al. 2006; Nalanagula et al. 2006).

By recording the eye movement behaviour of experts and showing the projected fixation position to other observers, recent studies have demonstrated that novices can detect more faults during aircraft fuselage inspection (Sadasivian et al. 2005) and during circuitry board inspection (Nalanagula et al. 2006). In addition, novices can benefit from such cues so as to make better clinical decisions (Jarodzka et al. 2012; Litchfield et al. 2010). For tasks such as problem solving and program debugging, which involve factors in addition to visual search, viewing another's gaze can result in shorter task completion times (Pomplun et al. 1996; Stein and Brennan 2004; Velichkovsky 1995). Of course, there are some situations where unambiguous verbal comments are generally preferred to another's gaze (e.g., Bard et al. 2007; Van Gog et al. 2009). As such, positive effects will likely depend on the task demands and the different ways in which eye movement patterns can be presented (cf. Jarodzka et al. 2012; Nalanagula et al. 2006). We return to some of these issues in a subsequent section below where we report some recent research that we conducted to examine these questions.

A final issue of importance concerns the fact that projecting gaze behaviour to observers involves "artificially" represented gaze (i.e., eye-movement patterns dynamically overlaid on a screen-based image of a task or problem), and therefore it is debatable whether the cognitive processes that are evoked are the same as those arising during normal gaze perception and gaze following. For example, the emotional expression of the person being watched is often taken into account when using normal gaze as a predictive cue (Bayliss et al. 2007) or when determining mutual gaze (Lobmaier et al.2008). Although this information is absent when using artificial gaze, this method does provide an opportunity to look at more complex gaze sequences rather than simple directional processes. Indeed, there have been reports of higher-order search strategies based on another's real-time gaze. For example, observers can regulate their own search behavior in a collaborative visual search task by strategically ignoring areas that they can see are being observed by their collaborator (Brennan et al. 2008). By taking advantage of these non-verbal gaze cues, observers have been shown to be able to reduce their search times significantly. Thus, even with artificially represented gaze, observers alter their behaviour based on the perceived processing of others.

Making Use of Attentional Hints and Gaze Cues in Insight Problem Solving

As outlined above, "insight" in problem solving arises when an individual who is stuck on a task and unable to make headway suddenly breaks free of their unhelpful thoughts and is able to find a solution. The two-string problem (Maier 1931), introduced earlier, is a classic example of an insight problem; another example is Duncker's (1945) radiation problem, where participants have to find a way to destroy a stomach tumour using lasers without harming surrounding healthy tissue

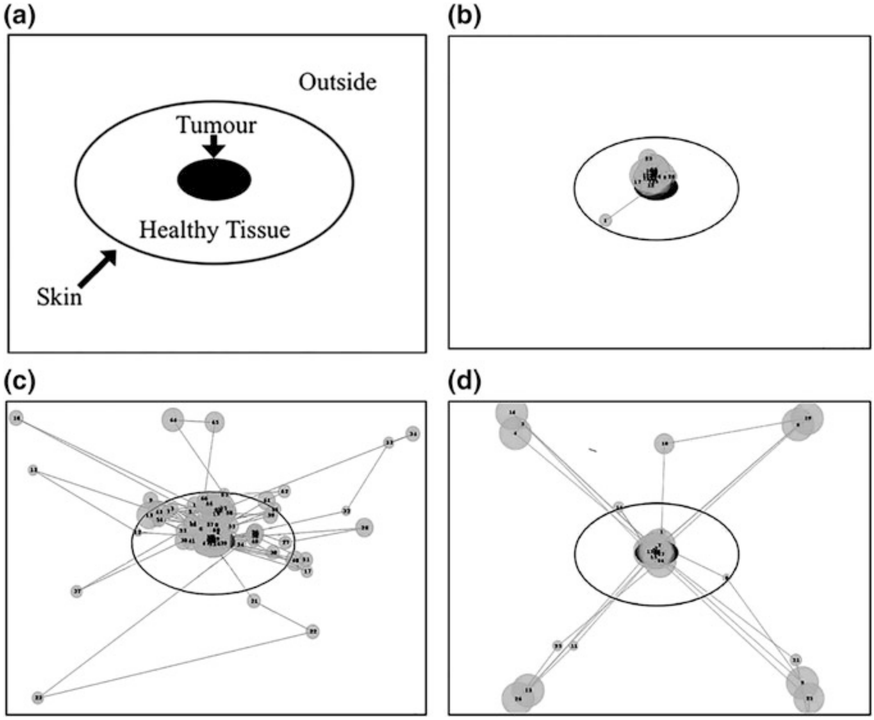


Fig. 6.1 Panel **a** shows a diagrammatic representation of Duncker's (1945) radiation problem, as used by Litchfield and Ball (2011), which was accompanied by a statement of the problem as follows: "Given a patient with an inoperable stomach tumour, and lasers which destroy organic tissue at sufficiently high intensity, how can one cure the patient with these lasers and, at the same time, avoid harming the healthy tissue that surrounds the tumour?" Panels **b**, **c** and **d** each depict a static scanpath that represents the 30 s of dynamic eye movement behaviour that was shown to participants in Litchfield and Ball's tumour-fixation condition (**b**), in their natural skin-crossing condition (**c**), and in their didactic skin-crossing condition (**d**). The dynamic scanpath that was presented to participants took the form of a moving blue gaze cursor

(see Fig. 6.1a). Although the solution involves only two critical components (i.e., converging multiple lasers of low intensity), problem solvers typically reach an impasse, and few solve the problem without hints such as visual analogies (Pedone et al 2001). Resolving the impasse requires some form of restructuring, with theories typically emphasising the role of unconscious constraint relaxation (Knoblich et al. 1999; Ohlsson 1992).

Recent research has examined the way in which eye movements and attentional guidance might influence solution success. Grant and Spivey (2003) demonstrated that participants who solve Duncker's (1945) radiation problem have specific fixation distributions just prior to solving the problem that are localised between the skin areas surrounding the tumour. They hypothesised that these fixation patterns may reflect the solver's mental imagery of where the lasers would have to be fired from to destroy the tumour. To test this hypothesis, Grant and Spivey conducted an experiment in which they made the diagram of the skin area more conspicuous by subtly increasing and decreasing its size. This “animated skin” condition engendered double the rate of solution success as seen in the control conditions, suggesting that implicitly guiding participants’ attention to the skin areas primed the perceptual-motor pattern associated with the solution. As they put it, “in-and-out eye movements themselves may have served as an embodied physical mechanism that jump-started a perceptual simulation (Barsalou 1999) of multiple incident rays, and wound up supporting the inference that multiple lasers could be fired (at low intensities) from different points outside the diagram” (p. 466).

Thomas and Lleras (2007) examined this perceptual-motor relationship further by sporadically guiding participants’ eye movements (via an unrelated number tracking task), either in a pattern related to the solution to Duncker’s radiation problem or in various unrelated patterns. Importantly, those participants who moved their eyes in a pattern related to the problem’s solution were more likely to solve the problem. Thomas and Lleras (2007) proposed that the improvement in solution rates arose neither because these cues increased the salience of the outer areas, nor because they increased the number of skin-crossing saccades. Instead, the improvement was a consequence of evoking a specific perceptual-motor pattern that embodied the solution to the problem, with the underlying shift of attention that accompanied the pattern being the crucial factor, rather than volitional eye movement per se. This interpretation was supported by Thomas and Lleras (2009a) in their follow-up research, which used the same paradigm as their earlier study (i.e., involving an unrelated number tracking task), but with participants this time having to do the tracking by shifting only their *attention* while keeping their eyes fixed on the centre of the display. This “attention-shift” condition produced similar solution facilitation to that which arose in the condition that permitted free eye movements during the tracking manipulation. Collectively, these results indicate that the attentional shifts that arise during scene examination can act as valuable, implicit hints that can guide how people think and reason.

The research that we have reviewed so far in this section demonstrates how eye movements and attentional shifts that derive from the *problem solver* can function to bootstrap their own efforts at insightful problem solving. But what about the situation where the problem solver has access to the eye movement patterns of another solver? We have recently examined this question—also in the context of participants tackling Duncker’s radiation problem (see Litchfield and Ball 2011). For this problem, it would be impractical to use a face-to-face model to convey the sequence of solution-specific eye movements described by Thomas and Lleras (2007). Instead, we decided to show the previously recorded eye movement patterns of a successful solver to observers so as to indicate where that individual looked during the task. As discussed earlier, although viewing “artificially” represented gaze is unlikely to involve identical cognitive processing to that which is deployed during normal gaze-following behaviour, viewing another’s eye movement patterns still allows observers to modify their information processing contingent upon the processing of others (e.g., Brennan et al. 2008; Nalanagula et al. 2006; Neider et al. 2010; Stein and Brennan 2004; Velichkovsky 1995).

A further issue that we were interested in addressing in our study concerned whether the model providing the eye movements was actually aware that other observers might be using their eye movements as visual cues. Gaze behaviour is often interactively "regulated" when people know that their eye movements are being observed during face-to-face situations (Kleinke 1986) or that their eye movements will be projected to others via eye-tracking equipment (Neider et al. 2010; Velichkovsky 1995). Indeed, the model providing the eye movements can deliberately *control* their gaze so that their intent is more effectively communicated to observers. There may, therefore, be a difference between viewing the eye movement patterns of a successful problem solver who is unaware that their eye movements are to be used as subsequent cues versus viewing another's eye movements as they actively try to convey the solution of the problem in a didactic manner (Velichkovsky 1995).

In our study (Litchfield and Ball 2011) we examined problem solving in three conditions (see Fig. 6.1). These conditions were: (1) "tumour-fixation", where participants viewed a problem solver's eye movement patterns focusing solely on the central tumour; (2) "natural skin-crossing", where participants viewed a problem solver's eye movement patterns *naturally* making skin-crossing saccades between the outside area and the central tumour from multiple directions; and (3) "didactic skin-crossing", where participants viewed eye movement patterns that *deliberately* made skin-crossing eye movements between the outside area and the central tumour from multiple directions. Performance in the tumour-fixation condition acted as a control, since this condition would not cue participants to make skin-crossing saccades associated with the solution.

As predicted, our results showed that participants who were encouraged to make skin-crossing saccades were more likely to solve the problem than those who simply focused on the tumour. Although the natural and didactic conditions led to equivalent final solution rates, participants in the didactic condition showed reduced solution latencies (i.e., faster insight solutions) relative to those in the natural condition. Overall, Litchfield and Ball (2011) conclude that there is good evidence to support the conclusion that viewing where another person looks can guide attention in a way that can positively affect problem solving.

Extending the Use of Gaze Cues to Real-World Problem Solving

One of the clear advantages of cueing visual attention by means of task-specific eye movement hints is the general applicability of this method to a wide range of real-world tasks. For example, in addition to studying how another's gaze can provide a problem solver with insight into how to treat a hypothetical tumour, we have also examined how the use of eye movement cues might facilitate problem solving performance in the domain of diagnostic radiography. Our particular research focus in this domain was on diagnosticians identifying pulmonary nodules (pre-cursors to lung cancer) in chest x-ray inspection (Litchfield et al. 2010). The eye movement patterns shown in this study were not designed to simulate a movement-based *thought process* relating to the solution to the problem (as in Litchfield and Ball 2011), but were instead presented so as to encourage observers to undergo the same series of *bodily movements* (i.e., eye movements) that experts make when examining chest x-rays and to focus on the same regions of interest that attract the attention of experts. In this way, these non-verbal, step-by-step hints served to influence observers' evaluative decisions regarding the presence of pulmonary nodules, and by doing so guided thought indirectly by first guiding attention.

In our study we used signal detection theory to assess diagnostic performance. In our initial experiment we found that both novice and experienced observers performed better at detecting lung nodules in an x-ray when previously shown the search behaviour of either a novice radiographer or an expert radiologist viewing the same visual display. In follow-up experiments we manipulated the task specificity of the presented gaze patterns (i.e., whether or not they related to lung nodule detection) as well as the perceived expertise of the model providing the gaze cues. These subsequent experiments established that only novices consistently improved their performance when shown an expert's search behaviour, and that these benefits only arose when the eye movements shown were related to a task-specific search. Moreover, a detailed examination of the contribution of model expertise indicated that even a naïve observer's search behaviour could help scaffold the decisions made by novices. Such evidence provides further support for the general applicability of gaze following behaviour and the pervasive functions associated with interpreting another's gaze (Emery 2000; Tomasello 1999). Our collective findings highlight that gaze following is a ubiquitous behaviour that can be used to support a variety of decisions and judgments in problem solving contexts. It is clear that knowing where another person looks can cue attention and shape thoughts and decisions, even when the problem solver is not in a face-to-face situation.

Conclusion

Over the past decade the theoretical conception of problem solving as being situated, embodied and interactive in nature has gained considerable momentum. One aspect of this changing view of problem solving has been the appreciation that as a highly social species, whenever humans are faced with a difficult task the most natural tendency is for them to exploit all externally available *hints* that provide clues as to how to proceed. Such hints may derive from looking (literally) to others for advice or from making use of other prompts that are available within the social and physical environments (Kirsh 2009; Kleinke 1986; Tomasello 1999; Sheperd 2010).

Our discussion in this chapter has taken much inspiration from Kirsh's (2009) recent sketch of a "theory of hints", which he presented as a first step toward a detailed consideration of what hints entail, where they stem from and how they provide a key basis for understanding how people overcome problems in concrete, real-world settings through a dynamic interaction with all elements of their extended environment. In contributing to Kirsh's theory of hints we have attempted to broaden the conception of hints to include the communicative cues that derive from other people's eye movements that can serve to direct attention and impart information in ways that can facilitate reasoning and problem solving. In reviewing a wide range of evidence for the role of eye movement cues as scaffolds for people's problem solving we have touched upon key notions in the emerging literature on grounded cognition (e.g., Barsalou 2008) and embodied cognition (e.g., Gibbs 2006; Spivey 2007).

There is no doubt that the role of eye movement cues in facilitating problem solving is complex and multi-faceted. Effects that have been observed include the embodied priming of motor sequences and perceptual simulations that enable problem solvers to enact the thought processes that underpin the solution to a problem (e.g., Litchfield and Ball 2011). Other effects involve implicitly cueing observers of eye movement patterns to imitate similar eye movement patterns when searching visual displays for possible pathology (e.g., Litchfield et al. 2010). This latter form of cueing brings with it a key *attentional* component in as much as imitative eye movements guide attention to regions of interest within the problem display, which, in turn, become subjected to inferential thought aimed at making a diagnostic decision. Eye movements can also function in many communicative contexts as an *explicit* cue to archive "joint attention", ensuring, for example, that people are attending to the same objects within the immediate environment and can thereby engage in shared reasoning about such objects (e.g., Becchio et al. 2008; Flom and Pick 2007; Hanna and Brennan 2007).

We conclude by suggesting that if cognitive science is serious about considering intelligence as an interaction between individuals and the hints and cues that derive from the physical and social environments (e.g., Barsalou 2010; Kingstone et al. 2003; Proffitt 2006; Sebanz et al. 2006; Wilson 2002), then considerable research time and effort will need to be devoted to developing the theories and methodologies outlined in this chapter. It is only by doing so that we will truly be able to quantify complex problem solving and reasoning behaviour from a situated, embodied and grounded perspective.

Acknowledgements This chapter is based on a paper presented at the Symposium on Distributed Cognition, University of Kingston, Kingston-upon-Thames, UK, held in July 2010, and sponsored and organised by Kingston Psychology and the Distributed Language Group. We are grateful for David Kirsh's invaluable feedback on this presentation, which inspired the connection between his sketch of a theory of hints and our interest in embodied and interactive eye movement cues in problem solving. We are also grateful to Stephen Cowley and Frédéric Vallee-Tourangeau for follow-up discussions on the work we presented at the symposium as well as for their critical comments on the present chapter.

References

Ball, L. J., & Ormerod, T. C. (2000a). Applying ethnography in the analysis and support of expertise in engineering design. *Design Studies, 21,* 403–421.

Ball, L. J., & Ormerod, T. C. (2000b). Putting ethnography to work: The case for a cognitive ethnography of design. *International Journal of Human-Computer Studies, 53,* 147–168.

Bard, E. G., Anderson, A. H., Chen, Y., Nicholson, H. B. M., Havard, C., & Dalzel-Job, S. (2007). Let's you do that: Sharing the cognitive burdens of dialogue. *Journal of Memory and Language, 57,* 616–641.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577–660.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645.

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science, 2,* 716–724.

Battersby, W. S., Teuber, H. L., & Bender, M. B. (1953). Problem solving behavior in men with frontal and occipital brain injuries. *Journal of Psychology, 35,* 329–351.

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition, 104,* 644–653.

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Sciences, 12,* 254–258.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition, 106,* 1465–1477.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8,* 535–543.

Busby, J. S. (2001). Practices in design concept selection as distributed cognition. *Cognition, Technology & Work, 3,* 150–160.

Butterworth, G., & Jarret, N. (1991). What minds have in common in space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9,* 55–72.

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception, 27,* 951–964.

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition, 29,* 1146–1152.

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61–83). Hove, England: Erlbaum.

Dominowski, R., & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg & J. Davidson (Eds.), *The nature of insight* (pp. 33–62). Cambridge, MA: MIT Press.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6,* 509–540.

Dunbar, K., & Blanchette, I. (2001). The in vivo/in vitro approach to cognition: The case of analogy. *Trends in Cognitive Sciences, 5,* 334–339.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58*(5, No. 270), i–113.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24,* 581–604.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211–245.

Flom, R., & Pick, A. D. (2007). Increasing specificity and the development of joint visual attention. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze-following: Its development and significance* (pp. 95–111). London: Erlbaum.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review, 23,* 523–552.

Gibbs, R. (2006). *Embodiment in cognitive science*. New York: Cambridge University Press.

Gibson, J., & Pick, A. (1963). Perception of another person's looking behavior. *American Journal of Psychology, 76,* 386–394.

Gilhooly, K. J., & Murphy, P. (2005). Differentiating insight from non-insight problems. *Thinking & Reasoning, 11,* 279–302.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14,* 462–466.

Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory & Language, 57,* 596–615.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (2010). Enaction, imagination, and insight. In J. Stewart, O. Gapenne, & E. A. Di Paolo (Eds.), *Enaction: Toward a new paradigm for cognitive science*. Cambridge, MA: MIT Press.

Jarodzka, H., Baslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika, B. (2012). Conveying clinical reasoning based on visual observation via eye movement modeling examples. *Instructional Science*, *40*, 813–827.

Kaplan, C., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22,* 374–419.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention researchers! It's time to take a look at the real world. *Current Directions in Psychological Science, 12,* 176–184.

Kirsh, D. (2009). Problem solving and situated cognition. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition*. Cambridge: Cambridge University Press.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin, 100,* 78–100.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory & Cognition, 25,* 1534–1555.

Kokinov, B., Hadjiilieva, K., & Yoveva, M. (1997). Is a hint always useful? In M. G. Shafto & P. Langley (Eds.), *Proceedings of the nineteenth annual conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Academic Radiology, 3,* 137–144.

Krupinski, E. A., Tillack, A. A., Richter, L., Henderson, J. T., Bhattacharyya, A. K., Scott, K. M., et al. (2006). Eye-movement study and human performance using telepathology virtual slides: Implications for medical education and differences with experience. *Human Pathology, 37,* 1543–1556.

Landrum, R. E. (1990). Maier's two-string problem revisited: Evidence for spontaneous transfer. *Psychological Reports, 67,* 1079–1088.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth a thousand words. *Cognitive Science, 11,* 65–99.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge, MA: Cambridge University Press.

Litchfield, D., Ball, L. J., Donovan, T., Manning, D. J., & Crawford, T. (2010). Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. *Journal of Experimental Psychology: Applied, 16,* 251–262.

Litchfield, D., & Ball, L. J. (2011). Using another's gaze as an explicit aid to insight problem solving. *Quarterly Journal of Experimental Psychology, 64,* 649–656.

Lobmaier, J. S., Fischer, M. H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology, 53,* 117–122.

Lobmaier, J. S., Tiddeman, B. P., & Perret, D. I. (2008). Emotional expression modulates perceived gaze direction. *Emotion, 8,* 573–577.

Maier, N. R. F. (1931). Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *Journal of Comparative & Physiological Psychology, 12,* 181–194.

Nalanagula, D., Greenstein, J. S., & Gramopadhye, A. K. (2006). Evaluation of the effect of feedforward training displays of search strategy on visual search performance. *International Journal of Industrial Ergonomics, 36,* 289–300.

Neider, M. B., Chen, X., Dickinson, C. A., Brennan, S. E., & Zelinsky, G. J. (2010). Coordinating spatial referencing using shared gaze. *Psychonomic Bulletin & Review, 17,* 718–724.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Ormerod, T. C., & Ball, L. J. (in press). Qualitative methods in cognitive psychology. In C. Willig & W. Stainton-Rogers (Eds.), *Handbook of qualitative research in psychology* (2nd ed.). London: Sage Publications Ltd.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking* (Vol. 1, pp. 1–44). London: Harverster Wheatsheaf.

Pedone, R., Hummel, J. E., & Holyoak, K. J. (2001). The use of diagrams in analogical problem solving. *Memory & Cognition, 29,* 214–221.

Pomplun, M., Ritter, H., & Velichkovsky, B. (1996). Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception, 25,* 931–948.

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science, 1,* 110–112.

Reid, F. J. M., Culverhouse, P. F., Jagodzinski, A. P., Parsons, R., & Burningham, C. (2000). The management of electronics engineering design teams: Linking tactics to changing conditions. *Design Studies, 21,* 75–97.

Sadasivian, D., Greenstein, J. S., Gramopadhye, A. K., & Duchowski, A. T. (2005). Use of eye movements as feedforward training for a synthetic aircraft inspection task. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2005* (pp. 289–300). New York, NY: ACM Press.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences, 10,* 70–76.

Sheperd, S. V. (2010). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience, 4,* 1–13.

Simon, H. A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge MA: MIT Press.

Smith, R. W., & Kounios, J. (1996). Sudden insight: All or none processing revealed by speed-accuracy decomposition. *Journal of Experimental Psychology; Learning, Memory & Cognition, 22,* 1443–1462.

Spivey, M. (2007). *The continuity of mind.* Oxford: Oxford University Press.

Stein, R., & Brennan, S. E. (2004). Another person's eye gaze as a cue in solving programming problems. *Proceedings of ICMI 2004: Sixth international conference on multimodal interfaces* (pp. 9–15). New York, NY: ACM Press.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge, MA: Cambridge University Press.

Symons, L. A., Lee, K., Cedrone, C. C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *The Journal of General Psychology, 131,* 451–469.

Tabachneck-Schijf, H. J. M., Leonardo, A. M., & Simon, H. A. (1997). CaMeRa: A computational model of multiple representations. *Cognitive Science, 21,* 305–350.

Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review, 14,* 663–668.

Thomas, L. E., & Lleras, A. (2009a). Covert shifts of attention function as an implicit aid to insight. *Cognition, 111,* 168–174.

Thomas, L. E., & Lleras, A. (2009b). Swinging into thought: Directed movement guides insight in problem solving. *Psychonomic Bulletin & Review, 16,* 719–723.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard, MA: Harvard University Press.

Van De Weil, M. W. J., Boshuizen, H. P. A., & Schmidt, H. G. (2000). Knowledge restructuring in expertise development: Evidence from pathophysiological representations of clinical cases by students and physicians. *European Journal of Cognitive Psychology, 12,* 323–356.

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior, 25,* 785–791.

Velichkovsky, B. M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics & Cognition, 3,* 199–222.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9,* 625–636.

Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes, 28,* 81–88.

Chapter 7
Cognition Beyond the Classical Information Processing Model: Cognitive Interactivity and the Systemic Thinking Model (SysTM)

Gaëlle Vallée-Tourangeau and Frédéric Vallée-Tourangeau

Abstract In this chapter, we propose a systemic model of thinking (SysTM) to account for higher cognitive operations such as how an agent makes inferences, solves problems and makes decisions. The SysTM model conceives thinking as a cognitive process that evolves in time and space and results in a new cognitive event (i.e., a new solution to a problem). This presupposes that such cognitive events are emerging from *cognitive interactivity*, which we define as the meshed network of reciprocal causations between an agent's mental processing and the transformative actions she applies to her immediate environment to achieve a cognitive result. To explain how cognitive interactivity results in cognitive events, SysTM builds upon the classical information processing model but breaks from the view that cognitive events result from a linear information processing path originating in the perception of a problem stimulus that is mentally processed to produce a cognitive event. Instead, SysTM holds that information processing in thinking evolves through a succession of *deductive* and *inductive processing loops*. Both loops give rise to transformative actions on the physical information layout, resulting in new perceptual inputs which inform the next processing loop. Such actions result from the enaction of mental action plans in deductive loops and from unplanned direct perception of action possibilities or affordances in inductive loops. To account for direct perception, we introduce the concept of an *affordance pool* to refer to a short term memory storage of action possibilities in working memory. We conclude by illustrating how SysTM can be used to derive new predictions and guide the study of cognitive interactivity in thinking.

G. Vallée-Tourangeau (✉)
Department of Management, Kingston Business School, Kingston University, Kingston, UK
e-mail: g.vallee-tourangeau@kingston.ac.uk

F. Vallée-Tourangeau (✉)
Department of Psychology, School of Social and Behavioural Sciences, Kingston University, Kingston, UK
e-mail: f.vallee-tourangeau@kingston.ac.uk

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_7

Some early psychologists took the position that only external, observable behaviour could be measured and hence inform models and theories of learning and behaviour. In its most radical presentation (e.g., Skinner 1935; Watson 1913) the behaviourist perspective rejected the idea that human beings are autonomous agents and instead cast them as biological organisms whose behaviour was entirely driven by external stimuli and measurable environmental contingencies. By way of consequence, the behaviourists' epistemological approach was to alter the environment to observe the consequence in behaviour and uncover stimulus-response relationships. The richness of human capabilities in general, however, soon made this position untenable. Some scholars retained the methodological tenets of behaviourism but rejected the idea that behavior was always purposeless and reducible to automatic responses to environmental cues (e.g., Tolman 1932, 1948). Others studied human memory (Bartlett 1995/1932), control systems (Craik 1948) or attention (Broadbent 1957) and pioneered a theoretical vocabulary that shaped the development of cognitive science in the second half of the 20th Century. In the words of George Miller, one of the early contributors to the so-called "cognitive revolution": "The grammatical rules that govern phrases and sentences are not behaviour" (Miller 2003, p. 141). In the work that followed, it was understood and accepted that the workings of the human mind could not be reduced to mere associations between events taking place "outside" the mind and behavioural responses in reaction to such events. Stimuli were recast as information inputs, and responses as the output of the mind, which itself took the centre stage as the processor of information. As cognitive psychologists sought to break in the mind's "black box" they effectively took over from behaviourists, with a strong implicit assumption about how cognition enfolds: namely, people receive information, process it, and produce a response such as an answer to a problem, a judgement, or a choice.

Somehow, however, as the behaviourist perspective was overthrown, with it went the proverbial baby. By focusing on the workings of the mind, sandwiched between the external stimulus and the resulting behaviour (Hurley 2001), cognitivists put a disproportionate emphasis on the head, with its mental processes, operations and computations at the expense of behaviour and its "situatedness". Cognition is studied in barren environments, where passive thinking subjects are conceived as mere consumers of information devoid of arms or hands. The issue with this approach is perhaps not immediately obvious: if a cognitive psychologist assumes that cognition is nothing but a mental process, then requiring her research participants to rely almost solely on their mind to carry out cognitive tasks seems reasonable enough. This state of affairs is further reinforced by the assumption that adult cognitive operations are ultimately and necessarily formal, involving operations on mental propositions rather than the messy physicality of the world (Inhelder and Piaget 2013). Effectively, while behaviourists reduced the workings of the mind associations between stimuli and responses, cognitivists are reducing them to mental operations. Both conceptions are wanting. Beyond adolescence, requiring anything other than your head to think is thought of as lazy, cheating, or at the very least appears to be a sign of poor cognitive functioning, mental disability or cognitive aging. By contrast, in the remainder of this chapter, we show that we could better understand how people actually think by allowing them to manipulate information both in their mind through mental processes and in their immediate environment through hands-on manipulations. Such manipulations, we contend, make a difference to the way people think. We present the Systemic Thinking Model (SysTM, adapted from Vallée-Tourangeau et al. 2015b), which builds upon the classical information processing model of cognition to account for this difference. We also introduce two new concepts: the concept of *cognitive interactivity* to refer to the emergence of cognitive events from the meshing of mental processing with the transformative actions of a thinking agent on her immediate environment; and the concept of *affordance pool* to refer to a short term storage of action possibilities in working memory that we conceive as sitting alongside the classical visuo-spatial sketchpad for imagery and the phonological loop for sounds. We conclude by illustrating how SysTM can be used to derive new predictions and guide the study of cognitive interactivity.

The Classical Information-Processing Model of Cognition

The classical information-processing model (Baddeley 2012; Wickens and Carswell 2012) represents cognition as originating from a series of input-processing-output events: We see or hear, we think, we respond. Figure 7.1 summarises this model (adapted from Baddeley 2012). According to this model, information flows from the environment to the mind as it is perceived and temporarily stored in one of two limited capacity systems: speech and sound information is stored in a phonological loop module through vocal or subvocal rehearsal whereas visuospatial information is stored in a visual-spatial sketchpad. While it is held in these modules, information is manipulated in an episodic buffer under the control of a central executive module. The central executive module is responsible for focusing attention, dividing attention between stimuli, and switching between tasks (Baddeley 1996, 2012). The episodic buffer is assumed to act as a buffer store, where a limited number of multidimensional information chunks are consciously combined from the phonological loop, the visuospatial sketchpad, and long term memory storage. This model assumes two “flows” of information. In the *bottom-up flow*, immediate stimuli present in the environment populate the phonological loop and the visuospatial sketchpad. In the *top-down flow*, information contained in semantic and episodic long term memory shapes the information stored, represented or processed. In all cases, information and mental processing all converge in the inner episodic buffer, from which the final response or cognitive outcome is produced. Taken together, these four modules make up the working memory system which is often cast as the cornerstone for complex cognition such as people’s ability to reason with novel information, also known as fluid intelligence (Kane et al. 2005).

This view of cognition and mental processing assumes a deductive flow of information in cognition where actions and behaviours (e.g., motor and verbal

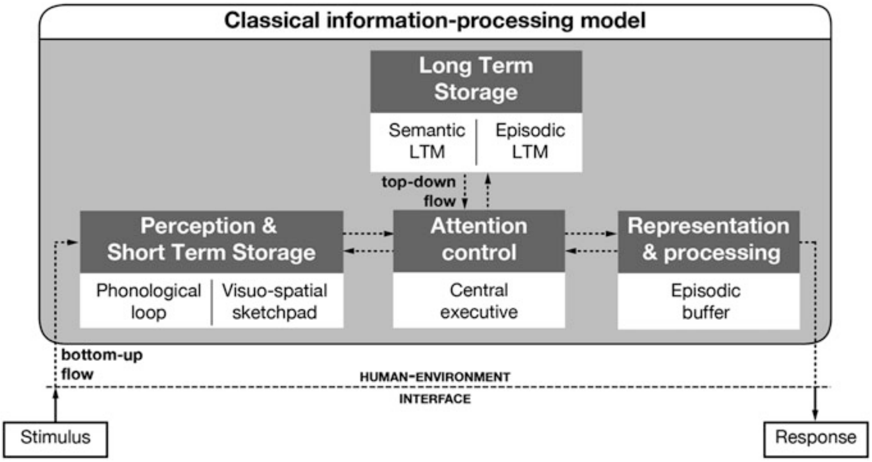


Fig. 7.1 The classical information-processing model (adapted from Baddeley 2012)

responses) are the end product of the application of intuitive or deliberative cognitive processes to the mental representation of an initial sensory input. This theoretical assumption is implicit in the methodology typically used to study thinking and decision-making. In the gambling paradigm, for example, a thinking agent will first perceive a stimulus, most likely in the form of a written text summarising two options: "Imagine you can choose between two lotteries: Lottery A gives you a 1% chance of winning $320 and a 99% chance of winning nothing whereas Lottery B gives you a sure gain of $3. Which lottery do you want to play?" (Adapted from Hertwig et al. 2004). According to Expected Utility Theory (EUT), a rational decision-maker should prefer the risky option (Lottery A) as its expected value ($3.2) is greater than that of the sure option ($3 in Lottery B). The classical information processing view describes the cognitive process by which human agents may reach their decision. This process begins with the mental representation of the different options in a more or less accurate fashion in the thinker's mind. For example, Kahneman and Tversky (1979) proposed that the preliminary cognitive processing of a choice problem will result in a simplified mental representation of the prospects, following editing operations such as rounding probabilities or outcomes. Next, thinking agents are assumed to apply evaluation operations to this internal representation as they mentally compute the weighted probabilities of gains and losses to compare the expected values of the prospects. Their final choice is assumed to be the option with the highest subjective expected value.

Within such a framework, good cognitive outcomes are attributed to adequate representations, accurate individual knowledge in long term memory, appropriate cognitive and attentional resources and motivational levels. Conversely, breakdown in performance is assumed to arise from inadequate mental representations, shortcomings in knowledge, cognitive busyness or depleted cognitive resources, and low motivational levels. Situational influences such as contextual aspects including the content of the problem (e.g., a gamble situated in a medical context where lives rather than money are at stake) or the perspective of the decision-maker (e.g., a patient deciding for himself or a doctor deciding for her patient) are acknowledged but only to the extent that they can affect the mental representation and mental processing of the problem (Wagenaar et al. 1988). Seemingly departing from this internalist conception of cognitive processing, some researchers have recently argued that although monetary gambles may share the properties of many real-world decisions, the way they are presented (as verbal descriptions stating all outcomes and their probabilities) may not reflect everyday life decisions where risks are not explicitly tabulated but instead need to be estimated from people's past experience. Deciding from experience rather than from description was found to affect final outcomes, as demonstrated by the robust description-experience gap effect (Hertwig and Erev 2009). Decisions from description are based on a written summary similar to the lottery example presented above. By contrast, decisions from experience are based on the sequential experience of uncertain outcomes. In an experiential setting, people first sample through each lottery outcomes instead of being presented with a descriptive summary of the lotteries. In Lottery A, they would experience winning nothing on 99% of the trials and winning $320 on 1% of the trials. In Lottery B, they would experience winning $3 on each trial. The description-experience gap results in a reversal of risk preferences. When choosing from a description, decision-makers prefer the risky option even if its outcome has a low probability of occurring, seemingly overweighting the probability of a rare but desirable event. Conversely, when choosing from experience, ceteris paribus, they prefer the safe option, thus seemingly underweighting the probability of a rare but desirable event in the risky prospect.

Despite the fact that thinkers appears more active in decisions from experience (in as much as they are in control of how much information they acquire), researchers still rely on the classical analysis of information processing to account for the description-experience gap. Thus, in decisions from experience, the mind is assumed to continuously update its representation of the probability value, experience after experience. The description-experience gap is thereafter assumed to orginate in a defective mental processing of probabilities informing the final choice in decisions from experience. Whether these representations are distorted due to the limited sampling of rare events, overweighting of late observations, or both, remains a debated issue (Hertwig et al. 2004; Hertwig and Erev 2009).

Besides representational issues, errors of judgement are also attributed to poor motivation to engage in effortful mental processing (Kool et al. 2010). Consider, for example, the bat and ball problem puzzle (Frederick 2005):

> A bat and ball cost $1.10.
>
> The bat costs one dollar more than the ball.
>
> How much does the ball cost?

This problem invites a quick answer (10¢) which is incorrect (the bat would cost $1 more—that's $1.10—and both would cost $1.20, not $1.10). The correct answer is 5¢. Kahneman (2012) writes “people who say 10¢ appear to be ardent followers of the law of least effort. People who avoid that answer appear to have active minds.” … “Failing these mini tests appears to be, at least to some extent, a matter of insufficient motivation, not trying hard enough … Those who avoid the sin of intellectual sloth could be called ‘engaged.’ They are more alert, more intellectually active, less willing to be satisfied with superficially attractive answers, more skeptical about their intuitions” (p. 45). To sum up, these examples illustrate how reasoning and decision-making performance is attributed to the characteristics of the thinking subject, the quality of her mental representation of the problem and the quality of her mental processing of this representation. In the next section, we propose an alternative perspective where performance is, instead, conceived as an emergent feature of a distributed cognitive system (Hollan et al. 2000).

The Classical Information-Processing Model Is an Inadequate Model of Thinking and Deciding

The classical information-processing model is a very compelling account which is supported by a variety of classic empirical phenomena. For example, people have more difficulty remembering series of items in the appropriate order when they “sound alike” as in *man, mad, can, cap, map* compared to series of items which have different speech sounds as in *pen, day, cow, bar, rig*; a phenomenon called the phonological similarity effect (Baddeley 1966). There is also evidence to show that people’s recall of digits presented visually can be hindered when they simultaneously have to ignore irrelevant speech sounds (Salamé and Baddeley 1986). Altogether, these findings point to the fact that memory traces are maintained through (subvocal) speech rehearsal. This mechanism is further evidenced by studies showing that when speech rehearsal is actively prevented through articulatory suppression (e.g., by asking participants to repeat “the” aloud while performing a task), the phonological similarity effect and the irrelevant speech effect disappear. In those instances, performance is simply diminished, whether or not items have similar speech sounds and whether or not irrelevant speech sounds can be heard in the background (Baddeley et al. 1975).

These processes impact cognitive outcomes beyond their impact on memory performance. For example, articulatory suppression is known to also impede mental arithmetic performance, especially when participants use counting strategies (Hecht 2002). Similarly, spatial reasoning ability is related to working memory visuospatial span, suggesting that the visuospatial sketchpad is involved in visuospatial reasoning tasks testing spatial visualisation. For example, in the dot-matrix task, participants are asked to remember a dot location in a grid while verifying the accuracy of equations. Performance on this task is strongly correlated with performance on the Paper Folding Test, a spatial visualisation test where participants are asked to mentally fold a piece of paper, imagine a hole punched through the folded paper, and infer what the unfolded paper would look like by selecting one of five alternatives (Miyake et al. 2001).

So far, these results suggest that mental processes (such as rote memorisation, mental arithmetic, or mental visuospatial reasoning) are depending on mental resources. However, the classical assumption that responses are inferred or deduced from mental processing is implicitly bounded by methodological individualism (Weber 1978/1922). This doctrine assumes that behaviour necessarily originates from the intentional states that motivate individual agents. This assumption in turn imposes a top-down hierarchy where individuals' thoughts precede their behaviour (Knappett 2005). If the individual agent is the ontological locus of cognition, then cognitive performance must be a reflection of the agent's mental capacities and abilities (Malafouris 2013). As we mentioned earlier, context is assumed to (e.g., Wagenaar et al. 1988) but only to the extent that different contexts may cue different representations and call attention to different information-processing strategies. Methodological individualism also constraints theoretical accounts which are putting a stronger emphasis on the role of the environment in thinking. Thus, the ecological rationality approach to cognition, for example, posits the cognitive machinery of the mind is best understood by examining how the mind exploits its immediate environment (Brighton and Todd 2009). As such, the ecological rationality approach falls well within the remits of the classical information processing model as it views cognition as emerging from mental computational mechanisms, albeit ones that have evolved to use so-called "natural" information inputs such as natural frequencies (e.g., Gigerenzer and Hoffrage 1995).

In line with the implicit methodological individualism that underpins the classical information-processing model, the research procedures commonly used to study complex cognition typically place participants in a passive information acquisition role where an experimenter controls the information they receive through auditory or visual channels. Experimenters then examine how a change in stimuli impacts reaction time or performance. Whereas many cognitive phenomena can and have been studied from this perspective, we contend that it offers a procrustean framework for studying complex cognition. In typical cognitive psychology experiments, stimuli are presented with paper-and-pencil questionnaires or computer display screens. This procedure allows experimenters to exert stringent control on how information is presented to participants. This strict control, however, obscures the potentially constitutive role of action and of the immediate environment in thinking and deciding. Since participants are, de facto, barred from handling, manipulating, and rearranging the information presented to them, methodological individualism reduces people to passive information processors who are modelled as if they typically remember, think, reason, solve problems and make decisions with their hands, as it were, tied behind their back and their eyes closed (see Hutchins 1995, Chap. 9 for a similar argument).

In other words, the classical information-processing model is adequate as long as one seeks to account for cognition arising from an information processing pathway where a unique final action, response, or behaviour is deduced from the processing of a mental representation. An important shortcoming of this model, however, is that it precludes the conception of the thinking process as evolving through a series of actions which will inform and transform a concurrent mental processing of the task information. The classical information-processing model also lacks a different kind of information processing loop, namely one where the next action, response, or behaviour is *induced from* the action possibilities offered by the immediate environment rather than *deduced from* mental processing. Yet, as we will discuss in the next section, there is accumulating evidence to show that thinkers sometimes act before they think when they can interact with their immediate environment. Far from being mere noise, those unplanned actions can transform mental processing and augment cognitive performance.

Rehabilitating the Hands: The Role of Cognitive Interactivity in Thinking and Deciding

People and things mutually influence each others. The key to overcome the vintage divide between the cognitivist and behaviourist conception of human behaviour is to move away from the mind-as-processor versus behaviour-as-reaction divide and instead focus on the nexus between mind and behaviour, which we call *cognitive interactivity*. Cognitive interactivity refers to the meshed network of reciprocal causations between an agent's mental processing and the transformative actions she applies to her immediate environment to achieve a cognitive result (see also Steffensen 2013; Steffensen et al. 2016). To embrace the central role of cognitive interactivity in cognition is to recognise that cognitive and behavioural outputs emerge, not from mental processing alone, but from the interweaving of mental and behavioural processes. The extent of such interactivity, however, depends on the affordances or action possibilities inherent in a given environment or situation. In interactive environments, thinkers become thinking *agents*, free to manipulate and fiddle with the information available in front of them. Interactivity results in physical changes in the environment in a way that best suits agents' thinking needs and flow. This, in turn, appears to facilitate information processing, and in many instances, to improve performance.

Different task settings will vary in the level of cognitive interactivity they foster. Cognitive interactivity may be nonexistent, limited to unobtrusive gestures, or unbounded. It is non-existent in the vast majority of experimental cognitive paradigms, either because agents do not feel entitled to impact their environment or because the environment is static and immutable (or both). In those instances, cognition is reduced to its barest form and limited to mental information processing. This type of cognition typically occurs in cognitive neuroscience research where participants are instructed not to scratch their head, swallow, open their mouth, yawn, inhale deeply, or shift their posture as any of these would produce spurious but irremediable noise in the neuroimaging data (e.g., see http://psychology.msu.edu/liulab/subj_info.html). Next, at the minimal level of interactivity, agents do not act to change their immediate environment, but they do act to support their own thinking. This is evidenced by the role of gesturing in memory retrieval and thinking (Goldin-Meadow 1999; Novack and Goldin-Meadow 2015). When free to gesture, individuals speak more fluently when their speech describes spatial elements, suggesting that gesturing can facilitate access to people's spatial lexicon (Rauscher et al. 1996). Gesturing can also lighten the cognitive load: individuals who were not allowed to gesture while explaining how they solved a mathematical problem also exhibited poorer recall on an interfering memory task compared to individuals who were allowed to gesture. This suggests that gesturing offloads some of the cognitive costs involved in the spoken task (Goldin-Meadow et al. 2001). In these instances, however, while gestures may offer a scaffolding support for thinking and deciding, they only represent a proto-level of cognitive interactivity.[1] The next level involves cognitive interactivity proper. When people can act upon and transform their immediate environment, their performance leaps up. A classic example is the finding that expert Tetris players begin rotating zoids so early (before 100 ms have elapsed) that they cannot reliably guess their shape (Kirsh and Maglio 1994). Such early rotations are also more frequent than what one would expect by chance alone so they are not random but they occur too early to be the enaction of a mental plan. Instead they contribute to performance by saving mental rotation effort and facilitating the identification of the zoid's type as well as the process of matching its contour with the existing mass (Kirsh and Maglio 1994).

Research from our own lab has also accumulated evidence that agents' performance improves if they can perform actions while attempting to solve a problem. For example, we presented participants with a Bayesian inference problem, a complex statistical reasoning task requiring them to revise the probability that a hypothesis is true, in the light of new data (Vallée-Tourangeau et al. 2015b). Half of the participants were presented with a standard pen-and-paper task where they could offload the content of their thought on paper but could not perform any action on the information content itself to support their thinking. The remaining half were presented with the same problem accompanied by a stack of playing cards, with each card representing some of the statistical information in the problem. In this setting, information was no longer frozen in a verbal description. It was distributed across the cards, thus offering participants the opportunity to mold their own external information layout as they progressed through the task. This always led to improved performance, above and beyond simply visualising the information through the cards: performance improved dramatically when participants were allowed to manipulate the cards and took it upon themselves to spend time actively changing the information layout by sorting and rearranging cards (Vallée-Tourangeau et al. 2015b). In other studies, we found that increasing manipulability of the environment facilitated insight (Vallée-Tourangeau et al. 2015a;

---

[1] A similar effect could be achieved through sub-vocalizing a string of articulatory moves. While this would be an instance of agentive activity, we would expect its ephemeral result to only offer a limited offloading support.

Weller et al. 2011; see also Fioratou and Cowley 2009), improved efficiency in mental arithmetic tasks (Vallée-Tourangeau 2013), facilitated learning in transformation problems (Guthrie et al. 2015), and enhanced word production in a Scrabble-like task (Vallée-Tourangeau and Wrightman 2010; see also Fleming and Maglio 2015; Maglio et al. 1999). These findings highlight the importance of cognitive interactivity: to understand cognition, that is, how cognitive events emerge, we need not only to understand how agents mentally process and represent problems but also how the actions they perform blend with their mental processes while they think the problem through.

SysTM: A Systemic Thinking Model of Cognition

The classical model of information-processing is ill-equipped to explain how higher cognitive results may emerge from cognitive interactivity. To account for, and further study, the role of cognitive interactivity on various cognitive operations such as mental arithmetic, insight problem-solving, or decision-making, we propose a dual-flow systemic model of cognition, where cognitive results arise from one of two processing loops: a deductive and an inductive loop (see Fig. 7.2). We characterize our thinking model as *systemic* to underscore our view that cognition emerges from a complex set of entities (human and non-human) that form an interconnected network of reciprocal causations. SysTM is intended as a framework for understanding and studying how information processing may be distributed across mental and material structures when an individual agent engages in a thinking task over a relatively short timeframe.[2] In SysTM, the human-environment interface separating the mental from the physical has been purposely removed to signal that the physical processing and the physical apparatus become an integral part of the cognitive substrate from which new thoughts and new actions may emerge.

The Deductive Processing Loop

In the *deductive processing loop*, the next action or response is deduced from the mental processing of a representation, akin to what is generally assumed in the classical information processing model (see Fig. 7.1). A key difference between this former model and SysTM is that, unlike the classical view, SysTM does not assume that this process follows a linear pathway from the initial stimulus to an

---

[2] As such, our use of the term ‘systemic thinking’ bears no conceptual resemblance to the term “systems thinking” coined by Senge (1991) in reference to the need of a shared vision and a focus on team learning to foster organizational transformation.

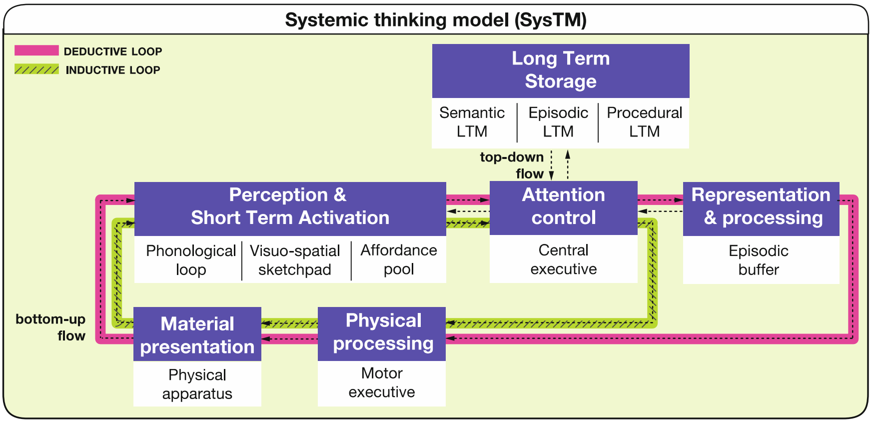


Fig. 7.2 The Systemic Thinking Model (SysTM)

intermediate stage of mental processing before reaching the final response. Instead SysTM assumes that processing evolves through a closed loop that includes both mental and physical processing of the information available. This looping implies that, rather than being temporarily stored, the perception of a stimulus *activates* sets of cues held in long-term memory (Ericsson and Kintsch 1995). Verbal cues becomes activated through the phonological loop and visuo-spatial cues, through the visuo-spatial sketchpad. SysTM also features an additional short-term activation component, the *affordance pool*, where motor-action sequences are temporarily activated. In a deductive loop, the activated cue sets shape a mental representation which is processed in the episodic buffer and ultimately directs the physical processing of the stimulus, provided that the environment is amenable to physical processing and that thinkers are empowered to act upon the affordances of their immediate environment. When engaged in a deductive processing loop, people execute planned motor actions on the material presentation of the stimulus. These actions are constitutive of thinking because they contribute to transforming the immediate perceptual field when they result in an alteration of the physical apparatus in their immediate environment. This altered material presentation offers new percepts to the mind, potentially attracting attention to new features of the informational landscape, thus reshaping mental representations and offering further opportunities for representation updating and mental processing, followed by further physical processing and so on.

The Inductive Processing Loop

SysTM does not assume that all information processing takes place in private thoughts but instead posits that some information processing can be offloaded onto the thinking agent's immediate environment when it features a manipulable information layout. In the deductive processing loop described above, such physical processing results from the execution of a mental plan formed through mental representing and processing. This processing pathway, however, is not sufficient to account for the emergence of cognitive results in interactive environments: not all actions need be carefully planned before they are executed. Agents may also act without a plan in their search for a fruitful physical information layout. To account for this type of external information processing, we propose that thinking agents' physical processing will at times be driven by a direct perception of action possibilities.

Several scholars have pointed out that stimuli may guide action in the absence of top-down (mental) categorization (Baber et al. 2014; Gibson 1979/1986; Greeno 1994; Norman 2002; Withagen and Chemero 2012). Some actions arise as individuals "follow materials" in a spatio-temporal trajectory (Ingold 2010; Vallée-Tourangeau and Vallée-Tourangeau 2014). Whereas the concept of direct perception has been proposed and defended before (Gibson 1979/1986; Wilson and Golonka 2013), it has yet to be integrated in a general cognitive framework. The doctrine that cognitive events may arise without mental processing of an inner representation is often summarily dismissed as a mere (behaviourist) heresy by cognitive psychologists interested in studying higher cognition (e.g., Fodor and Pylyshyn 1981; Gyr 1972) and cast as ill-suited to account for human higher cognitive processes (e.g., Shapiro 2011).

The issue, we contend, lies in the fact that proponents of the radical embodiement hypothesis often pitch direct perception as an alternative to traditional cognitive explanations and as an account of how perception works, in any situation (e.g., Wilson and Golonka 2013). It needs not be so. Freeing ourselves from the constraints of methodological individualism, we propose that actions are not always informed by mental processing and do not always result from a mental plan. Actions may also inform thinking before a plan is made or even before the informational content of a perceptual input is mentally represented and processed. The challenge in accounting for this possibility is to reconcile the view that perception nevertheless acts as informational input to the thinking agent while this input is neither mentally represented nor mentally processed.

The key to overcome this conceptual challenge lies in Norman's (2013) notion of "perceivable affordances". We conceive perceivable affordances as unmediated perceptions which inform the activities or actions that are possible within the individual's immediate environment. This concept is readily accommodated by the *affordance pool*, SysTM's third working memory component. Thus, perceived action possibilities temporarily activated in the affordance pool may compel the thinking agent to engage in an *inductive processing loop*, bypassing the need for representation and mental processing (see Fig. 7.2). People who are familiar with playing cards, for example, may immediately perceive that the cards afford picking up and sorting. They may engage in such physical processing without a mental plan of action. Similarly, the material presentation of the physical apparatus in the immediate environment may render a particular affordance highly salient for the thinking agent, in which case the affordance may be brought to light through the perceptual field in a bottom-up fashion. This implies that the direct perception of a given affordance may be influenced by past experiences. To account for this possibility, SysTM posits that motor-action sequences retrieved from procedural long-term memory may impact the affordances perceived in a given task, akin to the top-down mechanisms by which information retrieved from semantic and episodic long term memory might impact the information available in the phonological loop and visuospatial sketchpad components of working memory. Thus, when the thinking agent engages in an inductive processing loop, her actions are enacted via a direct path from the central executive to the motor executive, and eventually informed by information retrieved from procedural long-term memory (see Fig. 7.2).

Note that there is a conceptual ambiguity in the scope of affordances. For Gibson (1977, 1979/1986; see also Greeno 1994), affordances encompass all latent actions possibilities available in the environment, independently of the agent's ability to perceive them; we call these *latent affordances*. For Norman (2013), the critical affordances are those that are visible to the agent and affordances become visible in the presence of a perceptible sign or signal indicating what can be done; we call these *perceivable affordances*. In both conceptions, the ontology of affordances is unarguably relational: action possibilities depend on the relationship between physical or digital artefacts and agents' capabilities to interact with such artefacts. For example, a ball affords kicking as long as its size and weight are commensurate with the agent's ability to kick it. To account for what a thinking agent may do in a given setting, however, we need to focus on those affordances that become visible to the agent. Still, Norman's conception of perceivable affordances as depending only on the presence of appropriate signifiers in the immediate environment of the agent is wanting for our purpose. The process by which those signifiers are understood or translated as action possibilities remains under-specified. It conceals the potential top-down role of the agent's procedural knowledge, reflecting her behavioural repertoire, and what action possibilities may or may not be perceived. The affordance pool provides a route to address these issues by offering the means to specify the process by which affordances become visible to the agent.

Cognitive Interactivity

Taken together, the concepts of deductive and inductive processing loops provide a new framework to study how cognitive events may arise from cognitive interactivity where a thinking agent coordinates mental and physical resources to support her hypothesis testing, problem-solving or decision-making. As such, the systemic thinking model (SysTM) opens up a new agenda for research into higher cognition that transcends old debates: a key objective of the systemic approach is to understand how cognitive interactivity may produce cognitive events where the situated agent-environment ecosystem achieves a cognitive result (see also Steffensen et al.

2016). The systemic perspective presupposes that cognitive interactivity operates in a non-linear, but temporally situated fashion. Thinking and deciding need not to always follow a purely deductive processing pathway, as implied by the classical information processing model whereby perception precedes mental representation and mental processing which in turn directs physical processing. Conversely, thinking and deciding cannot be accounted for by a radical embodiment perspective where behaviour always emerges from the coupling of bodies with their specific environment without recourse to internal control structures including mental representation or mental planning (Wilson and Golonka 2013). Instead of pitching both frameworks against each other, our model reconciles these approaches by allowing either type of processing to take place through the spatio-temporal trajectory of cognitive interactivity. An agent may plan an action on her immediate environment before enacting it and the resulting change in the informational landscape may afford another action that does not require a plan. In other words, cognitive events may arise through mental or physical processing or both as the cognitive agent engages in a series of inductive and deductive loops in any given order.

Besides providing a new theoretical framework for understanding how cognitive results may emerge from information processing, the systemic thinking model (SysTM) can also serve as a guiding framework for research seeking to further our understanding of the complementary roles of mental aptitudes and environmental affordances in cognition. Mental aptitudes can be defined as the cognitive operations which are possible within the agent's mind. These may be stable (i.e., cognitive capacity such as working memory span, cognitive tools such as mental scripts and schemas) or transient (e.g., motivational and affective states). Mental aptitudes have long been studied by cognitive psychologists. As we have seen above, however, they have mostly been studied in dire settings stripped of most, if not all, environment affordances, thus reducing agents to thinkers paralysed by the dictates of methodological individualism, in the image of Rodin's *Le Penseur*. By contrast, the systemic thinking model (SysTM) highlights the need for studying how cognition emerges when agents can make full use of their mind and hands in settings rich of environmental affordances. Environmental affordances may be physical since actions possibilities depend on the manipulability of the physical informational layout, as well as social, since norms or vicarious influences may also constrain or foster action possibilities.

Different mental representations and mental aptitudes make different environmental affordances salient. Identified environmental affordances, in turn, govern the physical actions that the agent will implement to transform the material presentation of the information. This physical transformation will inform the mental re-representation of the task and guide future mental operations and physical actions, and so on. As Kirsh (2010, p. 441) puts it, such dynamic cognitive interactivity "allow us to think the previously unthinkable." Not only does cognitive interactivity saves memory and cognitive resources, but it also increases both the effectiveness and efficiency of cognitive processes. These processes become more effective as cognitive interactivity allows for more precise computations. They become more efficient as cognitive interactivity reduces errors and increases processing speed. In other words, adopting a systemic view of thinking calls for a re-assessment of the executive functions of working memory capacity such as attention allocation and switching. In a distributed cognitive system, executive functioning is no longer bound by the cognitive aptitudes of a thinking subject. Instead it is defined by an extended processing capacity that includes both the mental processing capabilities of a thinking agent and her physical processing abilities underpinned by the affordances she perceives in her immediate environment.

Key issues include better understanding of how environmental affordances are perceived and acted upon, both when behaviour results from careful mental planning and when it arises from direct perception. SysTM and a commitment to engineering thinking and deciding in interactive laboratory environments has the potential to usher in data that will cast a different light on models of problem solving [e.g., Ohlsson's (2011) redistribution theory; Weisberg's (2015) integrated framework], the role of working memory—and IQ—in thinking and deciding (e.g., Davidson 1995; Stanovich and West 1999), and on the cognitive abilities of the reasoners themselves (such as Bayesian reasoning, Vallée-Tourangeau et al. 2015a).

Qualitative analyses of agents solving problems in interactive environments along with verbal protocols and eye-tracking data may reveal the extent to which actions reflect the implementation of a plan and which don't. Likewise, much remains to be learnt about the relative proportions of different types of actions along different segments of the spatio-temporal trajectory that leads to a cognitive result. For example, Weisberg's (2015), Fleck and Weisberg (2013) recent integrated framework on insight problem solving is based on experimental data generated from a series of insight problems, some of which are presented with manipulable artefacts, some not. Participants' performance is substantially different with interactivity: restructuring is much more likely (Vallée-Tourangeau 2014). In turn, measures of working memory capacity in participants working on a difficult insight problem (the 17 animal problem, adapted from Metcalfe and Weibe 1987), predict none of the variance in participants' performance. Rather, the level of interactivity afforded by the problem environment alone explains whether participants can solve this problem (Vallée-Tourangeau et al. 2016).

One prediction from SysTM is that loading internal components of the agent's working memory would be particularly detrimental to task performance in conditions where physical processing is limited but not where the environment is rich of affordances. Mental arithmetic, for example, is known to involve the phonological loop to store intermediate values and the episodic buffer to carry out operations (Fürst and Hitch 2000). SysTM predicts that as the phonological loop is overloaded, the thinking agent may switch to an inductive processing pathway and store intermediate values in the physical world instead, provided that the environment affords such a switch.

We tested this prediction in a recent experiment. Participants completed a series of additions involving 11 single digit numbers. Although the task certainly does not exceed the mental arithmetic skills of numerate undergraduate participants, requiring them to complete it with their hands down neverthelsess resulted in systematic calculation errors. Errors rates are more dramatic when participants must also engage in articulatory suppression, repeating ‘the’ continuously as they work on the sums. However, increasing interactivity by presenting the task as sets of manipulable number tokens reduces both the impact of articulatory suppression and the magnitude of calculation errors (Vallée-Tourangeau et al. 2016). From the SysTM perspective, the working memory of the system configured by the coupling of an agent with numbered tokens and a flat surface to (re)arrange them, can better absorb the internal resource depletion caused by articulatory suppression. Equally interesting, independent measures of the participants’ level of mathematics anxiety revealed that the impact of suppression on calculation error was significantly moderated by maths anxiety—the higher the level of maths anxiety, the larger the impact of suppression—but only in the task condition that did not afford interactivity; in the condition where participants could manipulate the physical presentation of the sum, maths anxiety did not predict performance. This suggests that the cognitive resources of the system, rather than those of the agent alone, can augment arithmetic performance in the low interactivity condition, even among participants who are particularly prone to errors because of anxious thoughts about maths.

Future research could make use of SysTM to further explore the determinants of productive cognitive interactivity. For example, in our study of interactive Bayesian reasoning we found that only actions that involve a restructuration of the information subsequently promoted successful performance whereas actions that made minimal changes to the perceptual layout were ineffective in fostering a path to solution. Likewise, the characteristics of the affordance pool which we introduced to account for unplanned actions in thinking remain to be specified. We have found that loading the phonological loop impedes mental processing but not physical processing. It remains to be established whether loading the affordance pool (e.g., by asking people to press a pedal repeatedly while solving an interactive mental arithmetic task) would have a similar detrimental effect on performance in the presence of tokens.

Concluding Remarks

In this chapter, we reviewed the classical information processing model and argued that it offered a limited and limiting view of human cognition, bounded by methodological individualism. Next we reviewed empirical evidence pointing to the crucial role interactivity can play in explaining how individuals think. We proposed a new model of information processing—SysTM—aimed at addressing the shortcomings of the classical information processing model and to offer a framework for studying cognition that is free of the constraints of methodological individualism. At the core of SysTM lies the concept of cognitive interactivity, where cognition is conceived as emerging from the close coupling of mental aptitudes and environmental affordances. Finally, we derived specific predictions from SysTM to illustrate how this model could provide a guiding framework for studying cognitive interactivity in the future.

SysTM aims to account for higher cognitive operations including how an agent makes inferences, solves problems and makes decisions. It conceives thinking as a cognitive process that evolves in time and space and results in a new cognitive event (i.e., a new inference, a solution to a problem, a choice). As such, it provides a framework for studying one kind of distribution of cognitive process, namely that where mental and material structures are coordinated by an agent thinking on her own during a relatively short time episode. SysTM also offers a platform for overcoming the theoretical stalemate created by seemingly antagonistic views of cognition: namely the classical cognitivist view and the radical embodied view. Proponents of the former view argue that cognitive events emerge solely from mediated perception and computational processes that are executed mentally. Proponents of the radical embodied cognition view argue, instead, that cognition originates from direct perception and physical coordination. Instead of taking position for one of these qualitatively incompatible views of cognition, SysTM conceives mediated perception, direct perception, mental computations, and physical coordination of external information structures as different kind of processes, which all form part of the thinking agent's arsenal for addressing a cognitive task.

On a related note, in SysTM, attention control is conceived as a central executive process rather than a skill supporting the fine-grained sensorimotor coordination which can be observed in expert tool use (Baber et al. 2014). Again, these two conceptions of attention are not incompatible if we consider that expert tool use, on the one hand, and the type of cognitive tasks we have reviewed in this chapter, on the other hand, sit at the opposite end of a spectrum of cognitive activities people may engage in. At one end of the spectrum, we find learnt procedural routines and low levels of cognitive challenges. Skilled practice falls within that category. In expert tool use, attention will be primarily driven by the perception of affordances and SysTM posits that information processing will therefore primarily loop inductively by bypassing the mental representation and processing stage, and instead proceeds through a series of action-perception cycles. As mentioned earlier, the type of affordances cued in the affordance pool may be shaped by procedural long-term memory in a top-down fashion. So expertise could be reflected in a richer and perhaps more elaborate motor-action sequences readily activated, such as when the expert blacksmith engages in his or her craft (Wynn and Coolidge 2014). Simple cognitive tasks (e.g., "2 + 2 = ?") also fall within the category of learnt procedural routines that don't represent a cognitive challenge as such. Here, attention will be primarily driven by the perception of speech and sounds information via the phonological loop and SysTM posits this will be enough to retrieve the answer from long-term memory, thus also bypassing the mental representation and processing stage. For example, "2 + 2" is such a frequently encountered linguistic (verbal and written) expression that it no longer needs to be analysed in terms of its elements, but instead cues a direct association with another linguistic element, namely "4", that is reflexively produced like a conditioned response. At the other end of the spectrum, we find cognitive tasks that are both highly novel and highly challenging. Non-routine reasoning, problem-solving and deciding are typical examples of such cognitive tasks. When faced with such a task, SysTM posits that the thinking agent will engage in an "unplanned cognitive trajectory" (Steffensen 2013) that will involve a series of inductive and deductive processing loops. Whereas each cognitive trajectory may be unique, the balance between inductive and deductive processing loops should depend on the affordance landscape in the thinking agent's immediate environment (i.e., the Gibsonian latent affordances), the possible actions perceived by the agent (i.e., the Normanian perceivable affordances), her ability to process and plan her actions mentally, as well as her capacity to exert control on where to focus her attention.

The systemic thinking model (SysTM) substitutes methodological individualism for cognitive interactivism and cognitive interactivity is conceived as the core component of human cognition as it naturally unfolds, whether it is strategic or opportunistic in capitalising on fruitful but unpremeditated human-environment interactions. Under the cognitive interactivism assumption, thinking and decision-making are shaped by an interaction of inductive and deductive information processes that take place both internally, in the agent's mind, and externally, in her immediate environment: the material presentation of the initial informational landscape constrains people's representation of the task at hand. As mentioned before, SysTM was developped to model how higher cognition may be distributed across an agent and her immediate environment. In its current form, it is not fit to account for how cognition may be distributed across the members of a social group or across larger time episodes that would allow for the stabilization of knowledge and practice (Hutchins 2001), but future research may explore how it may be developed into a useful framework for studying those other kinds of distribued cognition as well. By putting cognitive interactivity at its core, SysTM not only offers the means to reconcile the vestigial chasm between a behaviourist approach (or the more recent radically embodied approach) and a cognitivist approach to understanding behaviour but also offers new avenues for investigating how people engage in higher cognitive processes.

References

Baber, C., Cengiz, T. G., & Parekh, M. (2014). Tool use as distributed cognition: How tools help, hinder and define manual skill. *Frontiers in Psychology, 5,* 1–14. doi:10.3389/fpsyg.2014.00116.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A, 49,* 5–28. doi:10.1080/713755608.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63,* 1–29. doi:10.1146/annurev-psych-120710-100422.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362–365. doi:10.1080/14640746608400055.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575–589. doi:10.1016/S0022-5371(75)80045-4.

Bartlett, F. C. (1995). *Remembering: A study in experimental and social psychology* (2nd ed.). Cambridge; New York: Cambridge University Press (Original work published 1932).

Brighton, H., & Todd, P. M. (2009). Situating rationality: Ecologically rational decision making with simple heuristics. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 322–346). Cambridge: Cambridge University Press.

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review, 64,* 205.

Craik, K. J. W. (1948). Theory of the human operator in control systems: II. Man as an element in a control system. *British Journal of Psychology: General Section, 38,* 142–148. doi:10.1111/j.2044-8295.1948.tb01149.x.

Davidson, J. E. (1995). The suddenness of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 125–155). Cambridge, MA: The MIT Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Fioratou, E., & Cowley, S. J. (2009). Insightful thinking: Cognitive dynamics and material artifacts. *Pragmatics & Cognition, 17,* 549–572. doi:10.1075/pc.17.3.04fio.

Fleck, J. I., & Weisberg, R. W. (2013). Insight versus analysis: Evidence for diverse methods in problem solving. *Journal of Cognitive Psychology, 25,* 436–463. doi:10.1080/20445911.2013.779248.

Fleming, M., & Maglio, P. P. (2015). How physical interaction helps performance in a scrabble-like task. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 716–721). Austin, TX: Cognitive Science Society.

Fodor, J. A., & Pylyshyn, Z. W. (1981). How direct is visual perception?: Some reflections on Gibson's "ecological approach". *Cognition, 9,* 139–196. doi:10.1016/0010-0277(81)90009-3.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives, 19,* 25–42.

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition, 28,* 774–782.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67–82). Hillsdale, NJ: Erlbaum.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ (Original work published 1979): Erlbaum.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684–704. doi:10.1037/0033-295X.102.4.684

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences, 3,* 419–429. doi:10.1016/S1364-6613(99)01397-2.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12,* 516–522. doi:10.1111/1467-9280.00395.

Greeno, J. G. (1994). Gibson's affordances. *Psychological Review, 101,* 336–342. doi:10.1037/0033-295X.101.2.336.

Guthrie, L. G., Vallée-Tourangeau, F., Vallée-Tourangeau, G., & Howard, C. (2015). Learning and interactivity in solving a transformation problem. *Memory & Cognition, 43,* 723–735. doi:10.3758/s13421-015-0504-8.

Gyr, J. W. (1972). Is a theory of direct visual perception adequate? *Psychological Bulletin, 77,* 246.

Hecht, S. A. (2002). Counting on working memory in simple arithmetic when counting is used for problem solving. *Memory & Cognition, 30,* 447–455. doi:10.3758/BF03194945.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15,* 534–539. doi:10.1111/j.0956-7976.2004.00715.x.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences, 13,* 517–523. doi:10.1016/j.tics.2009.09.004.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction, 7,* 174–196.

Hurley, S. (2001). Perception and action: Alternative views. *Synthese, 129,* 3–40.

Hutchins, E. (1995). Cultural cognition. *Cognition in the wild* (pp. 353–374). Cambridge, Mass: MIT Press.

Hutchins, E. (2001). Cognition, Distributed. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 2068–2072). Oxford: Pergamon. doi:10.1016/B0-08-043076-7/01636-3.

Ingold, T. (2010). The textility of making. *Cambridge Journal of Economics, 34,* 91–102. doi:10.1093/cje/bep042.

Inhelder, B., & Piaget, J. (2013). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*. Abington, UK: Routledge (Original work published 1958).

Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin Books.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–292.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131,* 66–71. doi:10.1037/0033-2909.131.1.66.

Kirsh, D. (2010). Thinking with external representations. *AI & Society, 25,* 441–454. doi:10.1007/s00146-010-0272-8.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18,* 513–549.

Knappett, C. (2005). *Thinking through material culture: An interdisciplinary perspective*. Philadelphia: University of Pennsylvania Press.

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139,* 665–682. doi:10.1037/a0020198.

Maglio, P., Matlock, T., Raphaely, D., Chernicky, B., & Kirsh, D. (1999). Interactive skill in scrabble. In *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society* (pp. 326–330).

Malafouris, L. (2013). *How things shape the mind: A theory of material engagement*. Cambridge, Massachusetts: MIT Press.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition, 15,* 238–246. doi:10.3758/BF03197722.

Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences, 7,* 141–144. doi:10.1016/S1364-6613(03)00029-9.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130,* 621–640. doi:10.1037/0096-3445.130.4.621.

Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral and Brain Sciences, 25*(01), 73–96.

Norman, D. A. (2013). *The design of everyday things* (2nd ed.). New York: Basic Books.

Novack, M., & Goldin-Meadow, S. (2015). Learning from gesture: How our hands change our minds. *Educational Psychology Review, 27,* 405–412. doi:10.1007/s10648-015-9325-3.

Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. Cambridge; New York: Cambridge University Press.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7,* 226–231. doi:10.1111/j.1467-9280.1996.tb00364.x.

Salamé, P., & Baddeley, A. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society, 24,* 263–265.

Shapiro, L. (2011). *Embodied cognition*. New York: Routledge.

Skinner, B. F. (1935). Two types of conditioned reflex and a pseudo type. *The Journal of General Psychology, 12,* 66–77. doi:10.1080/00221309.1935.9920088.

Stanovich, K. E., & West, R. F. (1999). Discrepancies between normative and descriptive models of decision making and the understanding/acceptance principle. *Cognitive Psychology, 38,* 349–385. doi:10.1006/cogp.1998.0700.

Senge, P. (1991). *The art and practice of the fifth discipline*. London: Century.

Steffensen, S. V. (2013). Human interactivity: Problem-solving, solution-probing and verbal patterns in the wild. In S. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (pp. 195–221). Dordrecht: Springer.

Steffensen, S. V., Vallée-Tourangeau, F., & Vallée-Tourangeau, G. (2016). Cognitive events in a problem-solving task: a qualitative method for investigating interactivity in the 17 Animals problem. *Journal of Cognitive Psychology, 28,* 79–105. doi:10.1080/20445911.2015.1095193.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Century.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55,* 189–208.

Vallée-Tourangeau, F. (2013). Interactivity, efficiency, and individual differences in mental arithmetic. *Experimental Psychology, 60,* 302–311. doi:10.1027/1618-3169/a000200.

Vallée-Tourangeau, F. (2014). Insight, interactivity and materiality. *Pragmatics & Cognition, 22* (1), 27–44. doi:10.1075/pc.22.1.02val.

Vallée-Tourangeau, F., Sirota, M., & Vallée-Tourangeau, G. (2016). Interactivity mitigates the impact of working memory depletion on mental arithmetic performance. *Cognitive Research: Principles and Implications*, *1*, 26. doi:10.1186/s41235-016-0027-2.

Vallée-Tourangeau, F., Steffensen, S, V., Vallée-Tourangeau, G., & Makri, A. (2015a). Insight and cognitive ecosystems. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 2457–2462). Austin, TX: Cognitive Science Society.

Vallée-Tourangeau, G., Abadie, M., & Vallée-Tourangeau, F. (2015b). Interactivity fosters Bayesian reasoning without instruction. *Journal of Experimental Psychology: General, 144,* 581–603. doi:10.1037/a0039161.

Vallée-Tourangeau, F., Steffensen, S. V., Vallée-Tourangeau, G., & Sirota, M. (2016). Insight with hands and things. *Acta Psychologica, 170,* 195–205. doi:10.1016/j.actpsy.2016.08.006

Vallée-Tourangeau, F., & Wrightman, M. (2010). Interactive skills and individual differences in a word production task. *AI & Society, 25,* 433–439. doi:10.1007/s00146-010-0270-x.

Vallée-Tourangeau, G., & Vallée-Tourangeau, F. (2014). The spatio-temporal dynamics of systemic thinking. *Cybernetics & Human Knowing, 21,* 113–127.

Wagenaar, W. A., Keren, G., & Lichtenstein, S. (1988). Islanders and hostages: Deep and surface structures of decision problems. *Acta Psychologica, 67,* 175–189. doi:10.1016/0001-6918(88)90012-1.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20,* 158. doi:10.1037/h0074428.

Weber, M. (1978). *Economy and society: An outline of interpretive sociology*. Berkeley: University of California Press (Original work published 1922).

Weisberg, R. W. (2015). Toward an integrated theory of insight in problem solving. *Thinking & Reasoning, 21,* 5–39. doi:10.1080/13546783.2014.886625.

Weller, A., Villejoubert, G., & Vallée-Tourangeau, F. (2011). Interactive insight problem solving. *Thinking & Reasoning, 17,* 424–439. doi:10.1080/13546783.2011.629081.

Wickens, C. D., & Carswell, C. M. (2012). Information processing. In G. Salvendi (Ed.), *Handbook of human factors and ergonomics* (4th ed., pp. 117–161). New York: Wiley.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00058.

Wynn, T., & Coolidge, F. L. (2014). Technical cognition, working memory and creativity. *Pragmatics & Cognition, 22,* 45–63. doi:10.1075/pc.22.1.03wyn.

Withagen, R., & Chemero, A. (2012). Affordances and classification: On the significance of a sidebar in James Gibson's last book. *Philosophical Psychology, 25,* 521–537. doi:10.1080/09515089.2011.579424.

Chapter 8
Time During Time: Multi-scalar Temporal Cognition

Jens Koed Madsen

Abstract Most models of cognitive function focus on internal structures while neglecting interactive and multi-scalar temporal elements. The current chapter advances the argument that cognition cannot be defined in isolation of three components: internal mechanisms (e.g. neural connections as well as our bodies), externally distributed interactions (e.g. sampling bias and other people), and multi-scalar temporal elements (e.g. socio-cultural background and habits). Further, the chapter advances the position that the definition of cognition necessarily is a functional concept. In this view, an objective world exists separate from cognitive functionality (and the individual performing the cognitive task). However, the derived functionality can only ever be an imperfect impression of this world to the extent that it is constructed and emergent as a result of the three aspects. This suggests a performed and multi-scalar temporal emergence of subjective selfhood that may nonetheless be modelled from a formal perspective. Finally, while most evidence in support of the multi-scalar temporal definition of cognition is indicative, the chapter briefly considers Agent-Based Models as a direct way to test plausible models of cognition that incorporates the three elements of cognitive functionality.

> Time flies over us, but leaves its shadow behind
> **Nathanial Hawthorne**

The Danish Council of Independent Research supports the research (ID: DFF – 1329-00021B).

J.K. Madsen (✉)
School of Geography and the Environment, University of Oxford, Oxford, England
e-mail: jens.madsen@ouce.ox.ac.uk

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_8

Introduction

Heraclitus famously proclaimed that a person could not possibly step twice into the same river, as the world is in flux, is dynamic, and is ever changing. This renders stability and permanence a futile dream. Against such a dynamic backdrop, humans engage in interactive, self-organised sense-making activities to behaviourally navigate the world. In doing so, humans rely on (constructed) memories, socio-cultural background, and our interaction with external artefacts and other people. Evidence suggests that these elements, perhaps causally connected with cognitive function, may be plastic, leaving fundamental cognitive capabilities such as beliefs, perception and memory malleable to situational change.[1] That is, cognitive capabilities may not be stable and exclusively internal, but rather malleable and interactive. This idea has entailments for our conceptualisation and operationalization of cognition, epistemology, and our sense of self. In this chapter, I advance the argument that the definition of cognition relies on elements necessarily and causally involved in generating the potential for emergent and malleable cognitive functions. Specifically, I argue that Multi-Scalar Temporalities (MSTs) is an integral and necessary aspect of cognition on par with neural networks and our bodies as well as the interacted context. In the MST cognition perspective, cognition not only goes beyond the brain and into the distributed context, but also goes beyond the particular moment in time in which we navigate the world. As such, MST cognition takes three aspects to be integral and necessary for cognitive function: internal mechanisms (e.g. neural connections as well as our bodies), externally distributed interactions (e.g. sampling bias and other people), and multi-scalar temporal elements (e.g. socio-cultural background and habits). In this view, there is an objective world separate from the cognitive functionality. However, the derived functionality can only ever be an imperfect impression of this world to the extent that it is constructed and emergent as a result of the three aspects. The derived may be skewed, which allows for cognitively reasonable processes that nonetheless bear little or no relation to the objective world. To explore emergent functionality, cognition cannot be isolated to any one of the aspects, but necessarily requires all.

Evidence, predominantly drawn from developmental and cognitive psychology, neuroscience, and language studies, provides indicative support for the 'self-illusion' hypothesis. This argues that the sense of a stable self or selfhood, as a phenomenological first-person experience, is illusory, as "[y]ou only exist as a pattern made up of all the other things in your life that shape you. If you take each away, 'you' would eventually cease to exist. This does not mean that you do not exist at all, but rather than you exist as the combination of all the others who complete your sense of self" (Hood 2012, p. 215). The chapter expands on this argument through the inclusion of MSTs. Crucially, the self-illusion does not

---

[1] 'Causal' refers to aspects that, if removed, would qualitatively change the outcome of a chain. Comparatively, supplementary aspects merely modulate the outcome of a chain. I am grateful to a reviewer who pointed out this terminological insufficiency.

remove agency, nor does it suggest that there is no delineation between the sense of self and the surrounding world (for example, if someone were to kill me now, my cognitive capabilities would quickly cease). However, it implies that definitions of cognition as an internal, stable, and operationalized sense of self that is located in and only in the brain is flawed from the onset and that a richer exploration of function requires a broader methodological foundation.

The chapter advances the position that cognition is internal and bodily, external and distributed, *and* temporally multi-scalar. It argues that these are all fundamental aspects of cognition, which shape how humans experience new situations. Take away any of these and cognitive functionality changes fundamentally. The chapter thus suggests a spatio-temporal foundation for cognition, which influence how researchers can model subjective, probabilistic reasoning paradigms that have been explored in cognitive psychology for the past couple of decades. Further, including temporal aspects as causal elements potentially allow for models to understand the ontogenesis of actions such as prejudiced violence, acts of terror, and radicalism, all of which appear extreme in isolation, but may manifest causally and gradually when expanding the temporal horizon. The MST cognitive perspective allows for emergent cognitive functionality from which derived impressions of the objective world feed into the cognitive mechanisms through which capabilities are enabled, shaped, and executed. As the emergent properties are attentionally, interactively, and subjectively derived, this allows for faulty learning processes that shape cognitive function across multi-scalar temporal scales, which further warrants the inclusion of MSTs in the definition and operationalization of cognition.

The chapter falls in four parts. First, we explore the internal and contextual influences on cognition and suggest that subjectively experienced information is malleable to such an extent that a definition of cognition needs to include external distributions. This is a key position in the Distributed Cognition paradigm. As externalisation and distribution are the subjects of other chapters in the book, they will be discussed only briefly. Second, we present different accounts of temporality, in particular Hemmingsen (2013) and the MST perspective (Uryu et al. 2014; Steffensen and Pedersen 2014). Third, although tentative and indicative in nature, we explore evidence relevant to the question of defining MST cognition. Finally, we consider consequences of the self-illusion hypothesis regarding emergent selves, probabilistic estimations, and language. For the current claim, we differentiate between necessary aspects and supplementary affordances for cognition. While changes in the latter would qualitatively change cognitive function, changes in the former may merely modulate cognitive function.

On the MST perspective presented here, cognition can be defined as a nascent, limited, and attention-directed property, which integrates elements from the brain, the body, situational constraints and Multi-Scalar Temporalities (MSTs). These properties are malleable and emergent, but are distributed in nature. This account agrees with the argument that a sense of a stable self and selfhood is illusory (Hood 2012) and that our definition of cognition needs to reflect this possibility. Going beyond the internal and the specific situation opens up for interesting empirical questions concerning the influence of MSTs on cognitive function typically studied in experimental and temporal faux-isolation.

The Brain and the Contextual

In order to develop a functional definition of MST cognition, we take point of departure from within the brain and extend the argument from there. This section reviews findings that suggest cognition is distributed and relational rather than solipsistic neural processes. The review suggests cognition needs to be defined as, at least, distributed between the internal and external at the same time, as both are necessary for emergence of cognitive potential.

Trivially, the brain is necessary for cognition. A definition of cognitive function could be internally oriented where the brain processes signals in order to solve behavioural tasks. This would operationalize cognition as a physical description of what goes on *within* the brain (connectivity, neural networks, etc.). There is a myriad of valuable work in neuroscience that shows the functional influence and necessary importance of the involvement of the brain in different tasks such as language comprehension and production (Scott 2012), decision-making (DeMartino et al. 2006), and other central cognitive functions. Without a brain we would certainly have no cognition.

Studies on cortical damage provide evidence concerning the necessity of neural networks and their influence on cognitive function. While cortical damage complicates and reduces cognitive competence, it does not remove the possibility of cognition entirely. It reduces rather than removes, as, despite significant neurological damage, a person may still be able to function cognitively in general despite reductions to task-specific functions. For example, a lesion in the left anterior insula has negative impacts on speech production, but does not necessarily impact speech perception (Dronkers 1996; Wise et al. 1999; Blank et al. 2002). Defining cognition as the possibility of cognitive function entails that the removal of *some* function does not entail the loss of cognition in general. Only removal of all cognitive functionality would entail removal of cognition in general. The fact that cognition can remain to a lesser degree is important for exploring evidence in which MST elements have been reduced or damaged. Importantly to the current argument, cortical damage studies show that reduction of a necessary function (neural damage) may still yield overall performance.

In the above, an initial definition of cognition suggested that studies should be internally oriented concerned with signal processing solving and predicting behavioural tasks. Numerous authors (some who contribute to this book) have argued that the definition and operationalization of cognition must go beyond the brain. Although differing in the specifics on how such an operationalization would occur, these authors argue that to perform cognitive tasks "…people use dispositions from both sides of the skull" (Valleée-Tourangeau and Cowley 2013, p. 2, see also Thibault 2011; Rączaszek-Leonardi 2014; Neumann 2014). On this view, the definition of cognition covers internal neural pathways *as well as* external interactive elements available within the context. Not merely affording cognition, the distributed perspective argues that external aspects are an integral and necessary part of the cognitive system, geared towards self-organising sense-making in a dynamic and interactive manner. On this view, external aspects are causally connected with cognition whereby the removal of such elements would reduce and qualitatively alter the possibility of cognition (much like aphasia reduces or qualitatively alters specific functionality without removing the potential for cognition entirely).

If we define cognition functionally, the distributed elements that are causally integrated within the cognitive system should be included in a broader definition of cognition on the same level as the internal neural pathways (although the involvement and weight of each element depends on the specific function). That is, going beyond the use of dispositions, people are influenced by their immediate choice context, which yields cognition to the extent that cognitive function might not be possible *without* these (or at least: would be significantly impaired or reduced). The following presents some evidence for the use and influence of external dispositions regarding communication and reasoning, both of which suggest that the definition of cognition needs to go beyond the brain.

Clark and Chalmers (1998, see also Hutchins 1995) argue that artefacts extend the mind such that external artefacts are part and parcel of the cognitive process. Hutchins exemplifies the perspective by considering the act of landing a plane. Rather than seeing such an act as an isolated neural activity, Hutchins argues that the pilot engages with a distributed system of bodies coordinating resources. The distributed perspective has been applied to language where "…human beings re-enact the cultural practices and patterns of the ecology" (Neumann and Cowley 2013, p. 15). Humans take, as it were, a language stance (Cowley 2011, see also Thibault 2011). Communication fundamentally exists as a dialogical system where no expression exists in a vacuum. Concurrently and in a similar vein, studies in communication and co-ordinated activities from the perspective of common ground and joint action suggest that communication relies fundamentally on the perception of the actions and beliefs of others in an inherently interactive manner (Garrod and Pickering 2004, 2009; Sebanz et al. 2006; Knoblich and Sebanz 2008; Pezzulo and Dindo 2011; Pickering and Garrod 2013; Rigoli and Spivey 2015).

Alongside communication, humans seem to align with the gaze and placement of bodies when solving complex problems. Pedersen (2012) shows how bodies align during medical consultations at a hospital in order to jointly solve a problem (see also Pedersen and Steffensen 2014 for a temporally oriented perspective on the medical case study), Steffensen (2013) shows similar patterns in a workspace, Vallée-Tourangeau and Villejoubert (2013) show participants using external affordances to solve complex reasoning problems. The evidence suggests that externally available interactions are integrally and causally connected with cognition to a degree where the removal of these would cause significant impairment to cognitive function. Akin to neurological damage, the removal of *some* of these elements would impair some cognitive function, but would still allow for cognition in general.

Pushing the notion of context and interactivity further, utility functions and decision-making are inherently influenced by interactions with the surrounding world. Although the following studies are not from the distributed cognition literature, studies in social and cognitive psychology, behavioural economics, and decision-making theory suggest that human beings do not enjoy stable utility functions in terms of probabilistic estimations of the traits of other (Madsen 2016), the framing of choices (Tversky and Kahneman 1981; McKenzie and Nelson 2003; McKenzie 2004), and the relative sample available for the decision (Stewart et al. 2006; Stewart and Simpson 2008). Rather, the frame, attention, and situation influences how humans functionally evaluate utility estimations of foundational properties. Collectively, the evidence suggests that information, inherently, is malleable to external manipulations and manifestations. Alternatively, one may argue that, rather than influencing probabilistic estimations, externalities are integral aspects of a self-organising system geared towards sense-making. If this is the case, we should expect different estimations given different internal make-up (e.g. two different people) or different externalities (e.g. different choice architecture). That is, the epistemic way in which humans deal with information and how we reason with that information is an emergent and interactive phenomenon. Thus, on a fundamental level, the way in which we perceive, engage with and use information is malleable given external differences.

Clark (1996) argues that the brain is a predictive engine that functions in tandem with the external world (see also Hohwy 2013 for a discussion of the brain as a predictive engine). These external artefacts and interactions are thought to be instrumental, integral, and essential elements of cognition such that the scope and a definition of cognition necessarily should be broadened to the extended, distributed sphere. If one defines cognition as the functional operations that make us reason, decide, and make sense of the world, it seems intuitively obvious that external properties and interaction with these should be included in the definition, as they are instrumental in shaping how we see information and choices, what decisions we make, and how we are able to understand our surrounding world. Indeed, it is difficult to grow a brain *without* the outside world to shape it.

In sum, there is evidence to assume that human perception of information, attention, cognitive involvement, and processing of language are all inherently interactive phenomena. This is in line with the distributed cognition view, which advances that, although trivially connected with the brain, human agency (and, for that matter, cognition) "…is constantly re-enacted as interactivity links us with the world (Neumann and Cowley 2013, p. 15). On this view, dispositions from both sides of the skull are causally connected with and therefore inherent to cognitive function and should therefore be included in a functional definition of cognition.

In the following two sections, I argue that the definition of cognition requires further extension. Not simply extended beyond the brain, evidence suggests that multi-scalar temporalities (MSTs) are as integral to cognitive function as the above evidence for interaction. I argue that theories (and models) of cognition need to integrate MSTs and, more fundamentally, that the definition of cognition needs to include MSTs. Further, I argue that MSTs, similar to internal and external aspects, is central for the epistemic foundation (potentially formalised as probabilistic estimations) of cognition. This is in line with Cowley and Steffensen (2015) who argue that humans are capable of ranging in time in order to manage affect, relationships, beliefs, and behaviours.

Temporality

Before considering indicative evidence for integrating temporality in cognition, we present ideas of temporality. There are different approaches to linearity of experienced time, memory (or the reconstruction of past events) as an element in cognitive processes. In this section, we discuss classic linear time, temporality as an affordance with cognitive events, and MSTs.

Classically, time can be considered from a linear perspective such that T_1, T_2, $T_{3,}$… T_n where T_n can only occur if and only if T_{n-1} has occurred. This fragments temporality into distinct and sequential units that may be separated from one another and represents a practical way of organising time in daily life (allowing amongst other things for appointments, train schedules and sequential measurements). On a practical level, this is a good organising tool for making reference to and use of temporal events, as it allows for the construction of sequential and causally connected accounts of events rather than an amalgamated and amorphous temporal blob of events. Although a useful tool to discriminate between causal relations of temporal events, this may not be a satisfactory conceptualisation from a cognitive perspective given the fact that we recall and make use of memories from different temporal events and amalgamate them in a cognitive moment, we integrate and embody different time-scales from biological mechanisms that have evolved across millennia to micro-scale events such as eye gaze, and the causal temporal order may be constructed and situationally emergent rather than stable and objective. Cognition, in other words, needs to move beyond classical linear temporality.

Temporal influences can be seen as the cognitive frame through which cognition can emerge. In this view, memories, socio-cultural background, historical events, language development and other temporal events are seen as sponsoring or affording cognitive potential or event in a given situation (see e.g. Hemmingsen 2013). In his chapter, Hemmingsen "… sets out to discover the effects of *Sponsorship* as an implicit experiential frame or even mentally thematized on *agency*" (p. 101)". Here, the temporal events are distinct and different from agency and cognition. Metaphorically, they are the stage upon which the play is performed. Discussing Gibsonian affordances, Hemmingsen (2013) argues that "Meaning is something, which is simply not "out there", independent of the *sense-making* agent (p. 100)". As will become evident, the current chapter agrees with the perspective that meaning is not 'out there', but rather arises as an emergent cognitive property or potential, which is negotiated in a distributed manner.

A slightly different account of temporal influences on cognition stems from Multi-Scalar Temporalities (see e.g. Uryu et al. 2014 expanded upon in Steffensen and Pedersen 2014). Rather than conceptualising time as a linear sequence, which can be accessed and which affords a frame through which cognition may emerge, the MST perspective argues that different temporal timescales are causally integrated *in* cognition. Steffensen and Pedersen (2014) suggest that this involves dynamic temporalities in which "*timescales* are invoked to emphasise that natural, living and social systems are self-organising wholes that cannot be captured by linear, uni-directional and reductionist models" (their emphasis, see also Enfield 2014).

Timescales acknowledge the influence and importance of elements that are temporally disjointed from or meshed within the situation in which the cognitive function emerges. They range from simple temporal aspects with relatively few causal frames to more complex temporal aspects with many causal frames. Additionally, they range from slow timescales such as evolutionary developments to very fast time frames such as unconscious eye gaze or phonemes in conversation.

Importantly, the MST framework argues that timescales are inherent and causally connected with the emergent property of cognition in the given moment. The framework is dynamic such that all timescales interact with the event in question in a causally connected manner. This is important, as the timescales involved are granted an active and causal rather than a passive and affording role. MSTs are part of cognitive processes in the same way that neural pathways are part of cognitive processes. Alongside a fundamental point regarding the inclusion of MSTs in cognition, Steffensen and Pedersen (2014) show how they might enrich interpretation of a communicative event in which references were made to historical events that go beyond the scope of the present situation.

In line with the distributed perspective presented in the above, the dynamics function both as a local and situated interaction as well as a non-local, trans-situational sociocultural framing (Steffensen and Cowley 2010). Further, the timescales in the MST are seen as causal to cognition. If this is true, we should expect detrimental effects on the cognitive system if one or more of these timescales were to be warped, damaged or removed from the system.[2] This suggests that the MSTs are not merely affording cognition; they are a causal *part* of the cognitive system. As MacWhinney (2005) argue, "all *mesh* in the current moment" (191). This distinguishes MST from an affordance account in the sense that timescales are not merely auxiliary means to support cognitive effort, but is *part of* the cognitive function.

It seems intuitive that MSTs should play a role in cognition in some fundamental manner. However, as humans are inherently caught in temporalities, it is difficult to collect data on the specific influence of MSTs on cognition. In the following section, however, we review some findings that suggest the constraining and/or

---

[2] As reviewed in section "Indicative evidence for MST cognition", there is a growing amount of evidence to support this claim.

enhancing effect MSTs have on cognition and ultimately advance the proposition that the definition of cognition should include MSTs as an intrinsic, inherent, and *causal* property of cognition.

Indicative Evidence for MST Cognition

If the removal of a temporal element causes significant decrease in cognitive function, it can be said to have an integral role in cognition and should therefore be included in the definition of cognition on par with internal and external aspects discussed in the above. On the other hand, some functional aspects may be supplementary affordances rather than necessary conditions. For example, losing a finger would not play an integral role in cognition. It may be distressing and cause dextral challenges, but it would not significantly impact cognitive capacity for understanding dextral capabilities. On this account, the finger is not an integral and necessary aspect of cognition, but an affording or auxiliary element. In the following, some evidence from language, socio-cultural, and memory studies is reviewed. The evidence suggests that MSTs are integral to cognitive function.

The Saphir-Whorf hypothesis states that cognition depends on language. This position can be expressed with different degrees of strength (ranging from deterministic to influential), but the main tenant states that cognition occurs, to greater or lesser degree, through the language the agents speak. Although some proponents have argued for a universal grammar inherent to brain structure (e.g. Chomsky 1980), Chater and Christiansen (2010) show that coordination-induction eases cognition function and argue that language shapes the brain rather than the brain shaping language (Chater and Christiansen 2008).[3] This constitutes language as a communal, dynamic, and interactive phenomenon rather than an entity emerging from within where two or more people share representations of their inner thoughts. As such, it is more aptly modelled as a probabilistic and coordinated phenomenon (see e.g. Chater and Manning 2006; Hsu et al. 2011) rather than a solipsistic transmission of mental states between detached individuals.

As mentioned in Sect. "The Brain and the Contextual", a number of studies provide evidence to suggest that communication is a sense-making and sense-saturated form of coordination between interlocutors in which meaning is a dynamic and emergent property rather than a transmission of mental states between two or more individuals. Further, if language is grounded in coordination between agents, it is worth recognising that language does not emerge with the emergent life of a particular individual. Language reaches beyond the confinements of the situation and the lives of the interlocutors in question. In the terminology of Thibault (2011, see also Cowley 2009), language is a first-order activity (action-perception)

---

[3] Presumably coordination-induction becomes easier over time with experience, suggesting a temporal relation in coordination aspects as well.

of a second-order entity (in interactive coordination with the available second-order language). As such MSTs have a profound causal influence on how language develops over time, the functionality of communication, and on the possible instantiations in any given communicative moment. If the Saphir-Whorf hypothesis has traction for cognition (whatever the strength of the hypothesis), the second-order, dynamic, and interactive nature of language entails that language as a centrally causal element of cognition cannot be traced to the individual, but is distributed across situation and MSTs.

Alongside the scope of language, there is evidence to suggest that cultural background influences how we see and make use of evidence. For example, cross-cultural studies show that participants from different cultures differ in responses to moral reasoning dilemmas (Tsui and Windsor 2001; Thorne and Saunders 2002), conditional reasoning tasks (Yama et al. 2010), how they integrate and make use of authority in reasoning tasks (Hornikx 2011), dialectical thinking (Zhang et al. 2015), as well as other elements of cognition.[4] Cross-cultural studies suggest that socio-cultural background is a causal part of how people see the world. In addition, behavioural habits influence how people engage with and use information (Gardner et al. 2011). Assuming the internal mechanisms and neural make-up of brains are somewhat similar across cultures, cultural differences cannot be accounted for by internal brain structures. Since culture itself is dynamic it will change over time in any social context. So, even the inclusion of cultural differences requires inclusion of temporal aspects. In such situations differences in cognitive function may more likely be ascribed to differences drawing on, in Steffensen and Pedersen's (2014) terminology, the human socio-culture timescale. Much like language, the way in which we engage with information draws on causal influences across many complex timescales.

Finally, we memory plays an integral part for cognitive functionality as detriments to these faculties entails loss of cognitive function and individual personality. Alzheimer's disease directly influences memory capabilities, and therefore offers fascinating evidence to support the hypothesis that cognition is mediated by and through multi-scalar temporal factors. A recent meta-review of evidence suggests that the disease influences Conscientiousness and Neuroticism[5] as level of neuroticism *increase* and conscientiousness *decrease* with the illness (Wahlin and Byrne 2011). Studies that corroborate these findings include Pocnet et al. (2011) and Jokela et al. (2014) who report lower openness, conscientiousness and extraversion and higher neuroticism while Lykou et al. (2013) report increasingly lower conscientiousness, extraversion, and agreeableness and higher neuroticism in patients with Alzheimer's compared with patients with mild cognitive impairments over a period of time. Consistently studies on Alzheimer's suggest that the illness

---

[4] Individual differences may account for some of this variance. The weight of culture and individual differences is a fascinating research topic that warrants more exploration.

[5] These are two of five general personality traits frequently used in the psychology of personality called OCEAN or 'the big five': Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (see e.g. Digman 1990; Johnson 2014).

can lead to changes in personality. This change is possibly due to degradation in memory, which allows for significant shifts in personality. Thus, the presumed stability of the self and cognition is affected by changes in memory structure, which provides indicative evidence that MSTs are causally and necessarily linked with cognitive function. As such, to borrow a phrase, memory can be seen as re-arranging a compost heap rather than recalling bits of inner film of pictures (Randall 2007).

Hemmingsen (2013, pp. 103–107) discusses a patient case (CW), which provides a case study of the importance of temporality in cognitive functionality (Hood 2012, pp. 54–56). He suffers from severe damage to the neural circuits that encode memory such that he is unable to form stable memories. This has left him severely cognitively impaired. Citing the memoirs of the wife, Hood describes CW as "… permanently trapped in the here and now" (see Wearing 2005) in a state where he experiences awakening as a conscious being for the first time in any given moment. However, CW is not completely without memories or causal structure. He recalls his wife and "… runs tearfully into her arms like a reunion of long-lost lovers when in reality she may have only left the room minutes earlier" (ibid), he is surprised to find chocolate in his hand (which suggests has access to a causal understanding of the world), he is proficient in English (and thus engages in second-order languaging), and is capable of some unconscious learning. In the above, it was posited that removal or damage to necessarily and integral elements in causal structures should cause significant detriments to functionality. The case of CW provides compelling evidence for the causal importance of memory in cognition and thus provides indicative evidence to suggest MSTs are causal and necessary elements of higher-order cognition.

MSTs appear to influence the language we speak, the way in which we reason, the way evidence is processed, personality traits, and cognitive performance in general (as seen in the case of CW). The evidence *suggests* the possibility that MSTs are causally involved with cognitive function rather than as an affordance device. As such, MSTs is an inherent property of the cognitive system in the same way that neural pathways and the contextual and interactive should be seen as causal mechanisms for cognition. Remove either of these and the cognitive system suffers detrimentally. On this view, cognition is an emergent, dynamic, and interactive flash in a multi-scalar temporal flux.

The above evidence is indicative and drawn from empirical studies on issues ranging from language comprehension to decision-making. Developments in model simulation have seen the emergence of Agent-Based Models (ABMs) to describe and predict behavioural and cognitive patterns that emerge when interactions between agents cause emergent states (i.e. states that are irreducible or unpredictable when isolating and extrapolating beliefs and behaviours from the individual agents in isolation). In general, ABM simulations involve three elements: agents, and environment and (cognitive and relational) rules (see Epstein and Axtell 1996). The dynamic aspect of ABMs show that time may potentially evolve markedly different belief and behavioural opportunities. Interactions and the growth of cultural differences across time may alter the potential for cognition and behaviour. As an example of the evolution of temporal nature of cognitive functionality, we may point to Agent_Zero (Epstein 2013). While different instantiations of Agent_Zero are developed for a broad range of phenomena, Epstein shows that aggregate behaviours can emerge from individual cognitive functionality that feeds of emotional and spatial interaction with other agents. He shows this is the case with riots, revolutions, and jury decision-making (Epstein 2013, part III). This allows for a individually based emergent cognitive function grounded in inner rules, interactive and spatial rules, and which requires temporal longitude for behavioural patterns to emerge. This further allows for a derived appreciation of the context in question that may or may not correspond with the objective world the agent inhabits. The derived becomes the foundation for the function. While the above evidence is concerned with empirical data, ABMs suggest that temporal influences may be modelled through interactions with the environment and other agents.

Although the literature on the role of MSTs in cognition is still sparse, the present evidence suggests that MSTs are integrated with principal elements of cognition. Given the relative recent suggestion of the MST perspective, however, the empirical and simulated evidence remains indicative and suggestive. The chapter therefore strongly encourages future research into the role of MSTs in cognitive function in order to explore the way in which they form a causal link in a cognitive model.

Consequences of MST Cognition

Throughout the chapter, I have advanced the argument that the definition of cognition should hinge on necessary and causally connected conditions for cognitive functionality.[6] Neural pathways are clearly such a condition. Removal of neural tissue fundamentally damages and impairs cognitive function. The argument was then advanced for the inclusion of situational and distributed factors. Finally, exploring evidence from different fields, it was suggested that MSTs causally contribute to the emergence of cognition and cognitive functionality. As such, a richer definition of cognition should include MST elements. This has significant consequences for conceptualising, describing, and predictively modelling cognition. In the following, we look at each of these.

Conceptualising (and defining) cognition in space, but without time is akin to exploring emergent and critical qualities of historical events without exploring socio-economic and cultural factors that shape the event. Without historical context,

---

[6] Both situational and temporal aspects go beyond the brain in fundamental ways that are not traceable or reducible to neural pathways. For example, emergence of group behaviour may not be reducible to he cognitive functions of the involved individuals, but draw on emergent properties that go beyond each individual. Agent-Based Models may be an ideal too to test the hypothesis that cognitive function in a larger network goes beyond the cognitive capabilities of each individual.

individual historical events become sudden, displaced, and may seem irrational or non-causal. They occur in a vacuum. For example, analysing the onset of the French revolution historians may consider past actions of the Ancient Regime, the levy of salt tax, power discrepancies of the three estates, and the different juridical practices in the various French provinces, and many other contributing factors. An analysis omitting these would appear deficient, without context, and lacking a fundamental level of understanding of the multiple causes and interactions that precipitated the situation. I posit that a conceptualisation and operationalization of cognition without temporal aspects will yield equally de-contextualised judgments that will not reflect the complex cognitive reality of vernacular activities. In particular, a given action may seem irrational and extreme in a temporal vacuum, but may appear causal, gradual and understandable when considered in a MST perspective.[7] This does not mean that all possible multi-scalar aspects of time have to be included in all empirical tasks (as this would be impossible), but rather entails the recognition that cognition, much like historical events, is defined and operationalized in the context of the MSTs that causally influence the possible scope for cognition and agency.

MST cognition opens for richer analyses of cognitive functionality such as reasoning, memory, and decision-making, if it includes neural pathways, interactive and distributed elements and MSTs. This allows for dynamic descriptions and analyses of the main elements involved in a particular cognitive task. For a neural-based (internal) definition of cognition, the relevant challenge in describing function is to determine the neural regions involved in the task, their connectivity, weight, individual function and so forth. Broadening the definition of cognition to include distributed and interactive phenomena as well as MSTs, this challenge remains, but is equally broadened. In addition, research should determine also the main causal connections for a given cognitive function, their individual weight, interaction of causally connected elements and so forth. This mirrors findings in neuroscience, which shows that some patches of neural networks are causally involved with some cognitive functions and not with others. We expect the same with MSTs. The fundamental challenge of describing and modelling elements causally connected with cognitive function therefore remains the same, but is more complex in line with the MST definition.

The fact that the MST cognition broadens the fundamental conceptualisation and allows for complex descriptions of cognitive functions naturally entails that computationally predictive models are equally expanded. Following descriptive changes as discussed in the previous section, formalised models of cognitive phenomena could take point of departure in the elements identified as relevantly and causally describing cognitive function. However, to generate predictive models, bottom-up

---

[7] Prejudiced violence is a good example of this. Isolated, such actions may appear irrational and unmotivated. MST analyses of such cognitive functions may provide an ontogenesis of hatred that cannot be located in the specific situation in which the hate-crime takes place. This broadens the epistemic field and allows for different causal connections that motivate violence (e.g. the upbringing of the person).

data-driven explorations would be needed to test and determine the extent to which the elements are explanatory, integrated, and ultimately predictive of the function. As with conceptualisation and description, broadening the definition of cognition from internal cognition to distributed MST cognition would not change the fundamental aims of modellers, but would enrich the potential elements that might be causally and significantly connected with the function in the model. As such, the MST cognition perspective entails a broadening of concept, description and formalisation, which integrates internal mechanisms (e.g. neural connections and cognitive effort), externally distributed interactions (e.g. sampling bias and other people), and multi-scalar temporal elements (e.g. socio-cultural background and habits). This acknowledges fundamental differences in world-views given substantial differences in any of the three main elements, which allow for dynamic models of cognitive function accounting for individual differences.

MST cognition has consequences for considering concepts such as the 'self', agency, and epistemology. Firstly, given the distributed and temporally scaled nature of cognition, MST cognition agrees with Hood (2012) that any notion of a stable unit of self is fallacious and illusory. Rather, the 'self' is an emergent property of consciousness called into being as a causal product of the three main elements: internal mechanisms, externally distributed interactions, and multi-scalar temporal elements (all of which are causally connected to the emergent cognitive function in question). Secondly, it suggests that any notion of complete self-determination and free will (understood here as the potential for complete cognitive independence) is also illusory, as cognition itself is caused by non-solipsistic factors. Evidence from judgment and decision-making suggests that our perception of utility functions are relative and malleable (Stewart et al. 2006; Vlaev et al. 2009), nudging theory suggests behaviour is malleable given different choice architecture (Thaler and Sunstein 2008), and priming suggests that perception of language (Rodd et al. 2013) and probabilistic estimations (Madsen, in press) are malleable and subject to change. To quote a phrase, we may believe ourselves to be Queens of our Castles or Lords of our Manors, but fundamental executive properties are not derived in a solipsistic manner. This does not entail epistemic or behavioural determinism, as the agent is capable of disentangling herself *to some degree* from either of these elements.[8] Finally, MST cognition suggests that our perception of the world is inherently subjective, malleable and non-solipsistic. If externally distributed interactions and MST elements are causally connected with emergent and dynamic 1st-order cognitive functions (and sense-making), fundamental epistemic properties of how humans see the world cannot be located exclusively within an individual. MST cognition therefore broadens the conceptualisation, description and modelling of cognition and points to fundamental epistemic questions.

---

[8] Note that internal mechanisms also provide cues that may be overridden. For example, craving ice cream while on a diet. Similarly, one's cultural background may provide strong cues for certain cognitive effort, which may (or may not) be overridden given reflection or determination.

Concluding Remarks

Cognition and cognitive functionalitycausally rely on more than brains in jars and more than brains in bodies. We extend outwards in a constant state of interaction with the external world. This relationship is fundamental to the function of cognition and thus should be included in the definition of cognition. In accordance with the Self Illusion hypothesis (Hood 2012), there is no stable 'you' or 'me', but rather an emergent entity. In this chapter, I have argued that Multi-Scalar Temporalities are part and parcel of cognitive function and therefore should be included in the *definition* of cognition. It has been argued that the contextual and the temporal are assistive devices *for* cognition (e.g. Hemmingsen 2013). In this perspective, the contextually and temporally influential are sponsoring cognition and perception. However, if the evidence advanced in the current chapter is to be taken seriously, I postulate that the contextual and multi-temporal, in the same way as neural pathways, are *part of* cognition (the weight, role, and interaction of each category and elements from each category depend on the cognitive function in question). As pointed out in Sect. The Brain and the Contextual, without a brain, we would certainly have no cognition. However, this is equally true for time (MSTs) and space (context). We witness the devastating effect of removing time when considering the above mentioned patient case of CW (and even here, while his memory faculties are severely damaged, he has a sense of time as it flows forward) while it is difficult to conceive of a cognitive opportunity with no space. These are such essential elements of cognition that we need at least three components for a theory of cognition: internal mechanisms (e.g. neural connections and cognitive effort), externally distributed interactions (e.g. sampling bias and other people), and multi-scalar temporal elements (e.g. socio-cultural background and habits).

If any neural tissue were removed, humans would still have some, albeit less effective, form of cognition. The same claim can be made for MST elements in cognition. For example, remove memory and you still have some, albeit less effective, form of cognition. This suggests that *neither* partial temporality nor partial neural pathways are sufficient conditions of cognition, but that cognition is an emergent property that stems from the neural, the body, the interactions, *and* the temporal. If we take this perspective seriously, we will have to classify *all* these properties as 'affordances' or, as suggested here, broaden the definition to include all properties when discussing the concept of cognition. This means that agency does not have a solipsistic foundation, but emerges as a distributed property. The malleability of utility and probability estimations suggest that this may be true. The mind, then, is a sense-making mechanism that allows for predictions of what may come, sense-making of why the organism is placed as it is (consciousness) and sense-making of how the present state came tobe. That is, future, present, and past. Crucially, this entails that we have no stable inner 'self' and that reflections upon the self are similar to reflections on the selves of others with the noted difference that we have access to more information about our emergent self, as we can feel our body, have more information about our own past than that of another, and that we have better access to desires, intentions, and internal considerations. However, this is a difference of quantity and quality of information rather than a difference in the type of reflection of inspection. As such, the self-self does not take an epistemologically privileged place.

The idea that MSTs influence cognition in a causal rather than an affording manner may be somewhat novel in a cognitive psychological perspective. However, other disciplines readily introduce multi-scalar temporal event when accounting for decision-making and cognition. For example, in his treatment of the Algerian war and its influence on the mental state of the inhabitants of Algeria, Fanon (1963) argues that "|i]ndividual experience, because it is national and because it is a link in the chain of national existence, ceases to be individual, limited and shrunken…" (p. 161). In this view, individual cognition *as purely solipsistic* does not exist. Literary explorations of cognition and the self also attest to MST as part of how people engage with the world. Cited in Madsen and Cowley (2014), the protagonist of Proust' *In Search of Lost Time*, Marcel, contemplates a church in Cumbria.

> I used to advance into the church… as into a fairy-haunted valley, where the rustic sees with amazement in rock, a tree, a pond… all this made of the church for me something entirely different from the rest of town: an edifice occupying, so to speak, a four-dimensional space – the name of the fourth being time – extending through the centuries its ancient nave, which, bay after bay, chapel after chapel, seemed to stretch across and conquer not merely a few yards of soil, but each successive epoch from which it emerged triumphant… (Proust 2001, p. 61)

To reiterate, this does not mean that there is no introspection or that we cannot discuss individual cognition as a functional entity, but it suggests that the very definition of cognition and subjectivity needs to include elements beyond neural networks in order to adequately describe and predictively model cognitive function and subjective, phenomenological epistemology. The subjective emerges from the distributed, the situational, and the multi-scalar temporal.

From an MST cognition perspective, findings that suggest the epistemic commitments of human beings are malleable and subject to change, that challenge the notion of a stable 'self' and a stable personality, and that shows the influence of memories, socio-cultural background, and choice frames do not delineate limitations of cognition or functionality. Rather, they are integral elements of what makes human cognition possible to begin with. In an interesting manner, it suggests that rationality and reasonableness can lead to different conclusions and decisions at different points in time, meaning that inconsistencies need not be irrational or 'wrong', but simply that the emergent subjective perception of the world to the self in that moment leads to different outcomes.[9] This opens up for exciting explorations regarding epistemology, decision-making, and other crucial cognitive phenomena, as it suggests a dynamic and fluctuating rather than a stable and universal foundation for cognition (i.e. the same elements may instantiate differently regarding weight, role, involvement and interaction in different individuals, situations, and

---

[9] Note, however, that this does not entail total ethical and normative relativism.

cultures). The possible descriptions and models that may follow from a broader notion of cognition should be explored in more detail to determine if the hypothesis is valid. At present the evidence is indicative, but I hope that the argument advanced throughout the chapter seems compelling: namely that reasonable cognition functionality causally depends on an interaction between internal mechanisms, externally distributed interactions, and multi-scalar temporal elements. The specific instantiation and function of these will depend on the function under investigation and the manifestation of that function at a specific place and time.[10] This points to exciting new programs of research that integrates MSTs in models of cognitive function.

References

Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., & Wise, R. J. (2002). Speech production: Wernicke, Broca and beyond, *Brain*, *125*, 1829–1838.

Chater, N., & Christiansen, M. H. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31,* 537–558.

Chater, N., & Christiansen, M. H. (2010). Language acquisition meets language evolution. *Cognitive Science, 34,* 1131–1157.

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing an acquisition. *Trends in Cognitive Sciences, 10*(7), 335–344.

Chomsky, N. (1980). *Rules and representations*. Oxford, England: Blackwell.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58,* 7–19.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Cowley, S. J. (2009). Distributed language and dynamics. *Pragmatics & Cognition, 17*(3), 495–507.

Cowley, S. J. (2011). Taking a language stance. *Ecological Psychology, 23,* 185–209.

Cowley, S. J., & Steffensen, S. V. (2015). Coordination in language: Temporality and time-ranging. *Interaction Studies, 16*(3), 474–494.

De Martino, D., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science, 313,* 684–687.

Digman, J. M. (1990). Personality structure: The emergence of the five-factor model. *Annual Review of Psychology, 41,* 417–440.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, 384, 159–161.

Enfield, N. P. (2014). Causal dynamics of language. In N. J. Enfield, P. Kockelman & J. Sidnell (Eds.), *Cambridge handbook for linguistic anthropology*. Chapter 13 (pp. 325–343), Cambridge: Cambridge University Press.

Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies: Social science from the bottom up*. Cambridge, Massachussetts: MIT Press.

Epstein, J. M. (2013). *Agenz_zero: Toward neurocognitive foundations for generative social science*. Princeton, New Jersey: Princeton University Press.

Fanon, F. (2001/1963). *The wretched of the earth*. London, UK: Penguin Modern Classics.

---

[10] The weight and influence of individual temporal nodes in such a system would be determined by careful examination of the case study and the data available for that system.

Gardner, B., Bruijn, G. J., & Lally, P. (2011). A systematic review and meta-analysis of applications of the self-report habit index to nutrition and physical activity behaviours. *Annals of behavioral medicine: A publication of the Society of Behavioral Medicine, 42*(2), 174–187.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8*(1), 8–11.

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science, 1*(2), 292–304.

Hemmingsen, B. H. (2013). The quick and the dead: On temporality and human agency. In S. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (pp. 93–112), Springer.

Hohwy, J. (2013). *The predictive mind*. Oxford, UK: Oxford University Press.

Hood, B. (2012). *The self illusion: Why there is no 'you' inside your head*. London, UK: Constable & Robinson Ltd.

Hornikx, J. (2011). Epistemic authority of professors and researchers: Differential perceptions by students form two cultural-educational systems. *Social Psychology of Education, 14,* 169–183.

Hsu, A. S., Chater, N., & Vitányi, P. M. (2011). The probabilistic analysis of language acquisition: Theoretical, computational, and experimental analysis. *Cognition, 120*(3), 380–390.

Hutchins, E. (1995). *Cognition in the wild*. London: The MIT Press.

Johnson, J. A. (2014). Measuring thirty facets of the five factor model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality, 51,* 78–89.

Jokela, M., Hakulinen, C., Singh-Manoux, A., & Kivimäki, M. (2014). Personality change associated with chronic diseases: Pooled analysis of four prospective cohort studies. *Psychological Medicine, 44*(12), 2629–2640.

Knoblich, G., & Sebanz, N. (2008). Evolving intentions for social interaction: From entrainment to joint action. *Philosophical Transactions of the Royal Society London B Biological Sciences, 363,* 2021–2031.

Lykou, E., Rankin, K. P., Chatziantaniou, L., Boulas, C., Papatriantafyllou, O., Tsaousis, I., et al. (2013). big 5 personality changes in Greek bvFTD. *AD, and MCI patients, Alzheimer Disease & Associated Disorders, 27*(3), 258–264.

MacWhinney, B. (2005). The emergence of linguistic form in time. *Connection Science, 17*(3–4), 191–211.

Madsen, J. K. (2016). The malleability and predictability of source credibility: American election candidates as a case study. In D. Mohammed & M. Lewinksy (Eds.), *Studies in argumentation and logic* (vol. 1, chapter 20, pp. 457–476).

Madsen, J. K., & Cowley, S. J. (2014). Living subjectivity: Time-scales, language, and the fluidity of self. *Cybernetics and Human Knowing, 21*(1/2), 11–22.

McKenzie, C. R. M. (2004). Framing effects in inference tasks and why they are normatively defensible. *Memory & Cognition, 32,* 874–885.

McKenzie, C. R. M., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulleting & Review, 10,* 596–602.

Neumann, M., & Cowley, S. J. (2013). Human agency and the resources of reason. In S. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (pp. 13–30). Dordrecht: Springer.

Neumann, M. (2014). Time scales of socio-cultural dynamics. *Cybernetics and Human Knowing, 21*(1/2), 66–79.

Pedersen, S. B. (2012). Interactivity in health care: Bodies, values and dynamics. *Language Sciences, 34*(5), 532–542.

Pedersen, S., & Steffensen, S. V. (2014). Temporal dynamics in medical visual systems. *Cybernetics and Human Knowing, 21*(1/2), 143–157.

Pezzulo, G., & Dindo, H. (2011). What should I do next? Using shared representations to solve interaction problems. *Experimental Brain Research, 211*(3), 613–630.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329–347.

Pocnet, C., Rossier, J., Antonietti, J.-P., & Gunten, Av. (2011). Personality changes in patients with beginning Alzheimer disease. *Canadian Journal of Psychiatry, 56*(7), 408–417.

Proust, M. (2001). *In search of lost time* (vol. I). London: Everyman's Library. UK: Random House.

Rączaszek-Leonardi, J. (2014). Multiple systems and multiple time scales of language dynamics: Coping with complexity. *Cybernetics and Human Knowing, 21*(1/2), 37–52.

Randall, W. L. (2007). From compost to computer: Rethinking our metaphors for memory. *Theory Psychology, 17,* 611–633.

Rigoli, L. & Spivey, M. J. (2015). Real-time language processing as embodied and embedded in joint action. In R. K. Mishra, N. Srinivasan & F. Huettig (Eds.), *Attention and vision in language processing* (pp. 3–23). Springer.

Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language, 68*(2), 180–198.

Scott, S. K. (2012). The neurobiology of speech perception and production—Can functional imaging tell us anything we did not already know? *Journal of Communication Disorders, 45,* 419–425.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Science, 10,* 70–76.

Steffensen, S. V. (2013). Human interactivity: Problem-solving, solution-probing and verbal patterns in the wild. In: S. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (pp. 195–221). Springer.

Steffensen, S. V., & Cowley, S. J. (2010). Signifying bodies and health: A non-local aftermath. In *Signifying Bodies: Biosemiosis, Interaction and Health, Catholic University of Portugal, Braga.*

Steffensen, S. V., & Pedersen, S. (2014). Temporal dynamics in human interaction. *Cybernetics and Human Knowing, 21*(1/2), 80–97.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology, 53,* 1–26.

Stewart, N., & Simpson, K. (2008). A decision-by-sampling account of decision under risk. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 261–276). Oxford: Oxford University Press.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth and happiness*. London, UK: Penguin Books.

Thibault, P. J. (2011). First-order languaging dynamics and second-order language: The distributed language view. *Ecological Psychology, 23*(3), 1–36.

Thorne, L., & Saunders, S. B. (2002). The socio-cultural embeddedness of individual's ethical reasoning in organizations (cross-cultural ethics). *Journal of Business Ethics, 35*(1), 1–14.

Tsui, J., & Windsor, C. (2001). Some cross-cultural evidence on ethical reasoning. *Journal of Business Ethics, 31*(2), 143–150.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211,* 453–458.

Uryu, M., Steffensen, S. V., & Kramsch, C. (2014). The ecology of intercultural interaction: Timescales, temporal ranges and identity dynamics. *Language Sciences*, *41*, Part A(0), 41–59.

Valleée-Tourangeau, F., & Cowley, S. (2013). Human thinking beyond the brain, in the wild. In S. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (pp. 1–14), Springer.

Vallée-Tourangeau, F., & Villejoubert, G. (2013). Naturalising problem solving, in the wild. In S. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (pp. 241–253), Springer.

Vlaev, I., Seymour, B., Dolan, R. J., & Chater, N. (2009). The price of pain and the value of suffering. *Psychological Science, 20*(3), 309–317.

Wahlin & Byrne. (2011). Personality changes in Alzheimer's disease: A systematic review. *International Journal of Geriatric Psychiatry, 26*(10), 1019–1029.

Wearing, D. (2005). Forever Today, Corgi

Wise, R. J. S., Greene, J., Büchel, C., & Scott, S. K. (1999). Brain systems for word perception and articulation, *The Lancet*, 353, 1057–1061.

Yama, H., Manktelow, K. I., Mercier, H., Henst, J-Bvd, Do, K. S., Kawasaki, Y., et al. (2010). A cross-cultural study of hindsight bias and conditional probabilistic reasoning. *Thinking & Reasoning, 16*(4), 346–371.

Zhang, B., Galbraith, N., Yama, H., Wang, L., & Manktelow, K. I. (2015). Dialectical thinking: A cross-cultural study of Japanese, and British students. *Journal of Cognitive Psychology, 27*(6), 771–779.

Chapter 9
Human Agency and the Resources of Reason

Martin Neumann and Stephen J. Cowley

Abstract The evidence shows that human primates become (relatively) rational actors. Using a distributed perspective, we identify aspects of human agency that raise important questions for sociocognitive science. In humans, we argue, agency does not centre on individual agents. Cognitive, social and linguistic outcomes depend on skills in moving in and out of aggregates that bind people, artifacts, language and institutions. While recognising the value of symbol processing models, these presuppose the embodied events of *human symbol grounding*. At a micro level, humans coordinate with others and the world to *self-construct* by cognising, talking and orienting to social affordances. We trace the necessary skills to sense-saturated coordination or *interactivity*. As a result of perceiving and acting on the environment, human individuals use the artificial to extend their natural powers. By using verbal patterns, artifacts and institutions, we become imperfect rational actors whose lives span the micro and the macro worlds.

Towards Sociocognitive Science

Humans make extensive use of cultural resources that include languages. Alone of the primates, their infants develop in a normative world that allows species to develop a distinctive type of agency. This is manifest individually and, just as strikingly, in how human groups implement long-term plans. For historical reasons, attempts to explain human agency have typically sought its origins in, not a history of acting in the world, but genes and/or brains. While biological factors are crucial, we argue that their function is to allow living beings exploit the environment to develop cognitive, social and linguistic skills. The process arises as infants learn to

---

M. Neumann (✉)
Jacobs University Bremen, Bremen, Germany
e-mail: ma.neumann@jacobs-university.de

S.J. Cowley
Centre for Human Interactivity and the COMAC cluster, Slagelse, Denmark
e-mail: cowley@sdu.dk

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_9

talk or, in exceptional cases, make use of signing. We therefore argue that it is through participating in language-activity that human primates *become* rational actors. Ours is thus a reworking of what makes us human in terms of Durkheim's (1895) famous claim that only the social can ground the social.

Embodied coordination enables infants to self-construct as members of a culturally extended ecology. Once we recognise that social predispositions for embodied coordination are functionally reorganized by encultured body-world activity, Durkheim's view ceases to appear circular. Social behaviour arises as we are moved by others to coordinate in increasingly complex ways. By hypothesis, bodies self-organize by learning to cognise, speak and act strategically: human agency develops within aggregates that bind together artifacts, institutions, and ways of experiencing the world. It is a human capacity for moving in and out of such aggregates—for exploiting embodied coordination—that enables social action to drive the self-construction of rational human actors. In tracing human agency to cognition beyond the brain, a history of coordination is seen as the basis of knowledge. Infants use spontaneous activity in learning to orient to a population's linguistic and other social practices. On the distributed perspective, this shapes human agency—infants discover ways of achieving effects that are, inseparably, cognitive, linguistic and sociological.

The Distributed Perspective

Human agency has previously been traced to how acting and thinking are honed by the demands of sociocultural environments. This is done in, for example, activity theory, cultural psychology and the pragmatism of Mead. What is novel to the distributed perspective is the view that agency results from acting and perceiving in socially distributed systems. For readers of *Beyond the Brain*, the idea will be familiar. The agency of a pilot who lands a plane is non-local in that, as Hutchins (1995) shows, it is distributed across bodies (and brains) that coordinate an aggregate of resources. The pilot uses readings from the instrument panel, messages from ground control, and real time interaction with a co-pilot. Far from centering on a neural system, agency arises in *acting* with material, cultural, and social structures. Ethnography makes that clear. To understand how such aggregates function, however, systemic output must be separated from construals and associated actions. In Tetris, human-computer aggregates rely on both and epistemic actions that use the world to simplify cognitive tasks (Kirsh and Maglio 1994). Epistemic actions depend on sense-saturated coordination or player-computer *interactivity*. They change on-screen resources in ways that suit the player's expert knowledge. In conversations, tight coupling allows people to concert by using, not just what is said, but also how speech and movements are integrated in rapid or pico time-scales.[1] In tight coupling that uses cultural resources, in Cronin's (2004) terms, we are cognitively cantilevered into the Umwelt. Coordination that is faster than conscious perception drives spontaneity by linking expertise with the results of joint action. The thesis of our paper is that this sense-saturated interactivity shapes human agency by giving us skills in dealing with artifacts, people, and languages.

In ethnography, language is identified with the words that are actually spoken (and which can be transcribed). However, its functionality depends on the fact that language too is *distributed* (see, Cowley 2007c, 2011a; Thibault 2011). In its paradigmatic face-to-face settings, language contributes to action by social aggregates: it is activity in which human beings re-enact the cultural practices and patterns of the ecology. Since it is both situated and verbal, people draw on each other's thinking. This is possible because, unlike animal signalling, languaging has a non-local aspect. Embodied activity links circumstances to past situations by virtue of how we perceive verbal patterns. During languaging, cultural constraints prompt real-time construal. In conversations, verbal patterns arise as concerted acts of articulation (or signing) are accompanied by facial and other expression. Since the results are both embodied and non-local, infants can link metabolic dynamics with between-people coupling. Without hearing what is said, interactivity—and the feeling of thinking—sensitise them to normative resources. While initially reliant on circumstances, they gradually come to hear language as consisting in thing-like parts or to *take a language stance* (Cowley 2011b). Once utterances sound like utterances of something, perceived results or *wordings* can be used in, for example, asking about things. Language, agency and rationality are, irreducibly, individual *and* collective. Cognition links the world in the head with the physical, linguistic and cultural processes of an extended ecology (Steffensen, 2011)—a place where individual actions carry cultural value. Human agency self-organizes during individual development. Infants use circumstances to become intentional and, later, make use of the resources of reason. Since language is ecological, dialogical and (partly) non-local, rationality co-emerges with individual agency. Although based in interactivity, ways of feeling out the world are supplemented by intelligent—partly conformist—use of external resources.

Agency and Human Agency

The term agency can be applied to people, animals, social institutions and inorganic processes. In the first place, it is therefore useful to distinguish physico-chemical agents and living systems. Only in biology do systems (and lineages of systems) set parameters (see, Pattee 1969, 1997) that allow them to measure and control aspects

---

[1] This is the time-scale within which gestures are made and syllables articulated. It can broadly be associated with a window of around 200 ms. It is especially important in prosody (see, Cowley, 2009) but, at this scale, interactivity is full-bodied (for detailed discussion, see Thibault, 2011).

of their environments. Living systems are adaptive and yet also able to maintain autonomy (see, Di Paolo 2005). Adaptive self-organization allows even single celled bacteria to explore their worlds by linking genes and metabolism with viable use of available resources. Flexibility increases in living-systems that use brains, development and learning. Yet, these processes too depend on self-organization or how organic systems exploit the world beyond the body (including other organisms; Thompson 2007). Organisms are aggregated systems whose parameters link a lineage's history with, in embrained species, experiential encounters with the world.

In evolutionary time, organisms show flexibility as they adapt to the world and, more strikingly, adapt the world to them. For Jarvilehto (1998; Järvilehto et al. 2009), the world of the living is to be conceptualised in terms of interlocking Organism-Environment Systems (O-E systems). The necessity of the view appears with perceptual learning: as Raguso and Willis (2002) show, foraging hawk-moths, learn about local sets of flora. In optimising their behaviour, their partly plastic brains link genetically based self-organisation with learning about an environment. Further, as conditions change, they alter the parameters. The example is apposite in that, while such intelligence is organism-centred, this does not apply to all insects of comparable neural dimensions. In bees and other eusocial insects, cognition serves the colony rather than individuals. Below we argue that languages, technology and money make humans partly eusocial. Since we live in a culturally extended ecology, we are *hypersocial* beings (Ross 2007) whose primate intelligence is extended as we orient to eusocial resources that function as cultural (second-order) constraints. Individual-environment relations thus transform individual experience, learning and ontogeny.

Human uniqueness depends on neither our hypersocial nature nor our propensity to exploit the world beyond the brain. What is most striking about human agency is how it combines artificial rigidity with our natural flexibility. As organism-environment systems, we detect *rationality*; as populations, we tend to act in line with utility calculations. Uniquely, humans are partly biological and partly rational. As individuals, we grasp rules (imperfectly), ascribe minds to agents, plan, take part in social institutions and make use of wordings, tools and machines. However, we draw on the resources and skills of populations. How is this to be explained? While bound up with learning to talk (not to mention literacy and numeracy), human agency also uses artifacts and institutions. These perform a dual function both as boundary conditions and as flexible constraints: they serve to measure and also to control. Given the relative predictablity of wordings, we extend our natural intelligence. Accordingly, we now turn to how coordination alters a social actor's sensorimotor and cultural strategies. Coming to act in line with utility calculation depends on learning to concert movements with those of others, exploit available social strategies and, using these, gaining skills in using the artifacts and cultural resources of a community.

Language and Languaging

Since the eighteenth century human nature has been associated with a mental locus of *ideas*. On such a view, language becomes a transparent conduit between minds (Locke 1690) used to construe verbally-based thoughts (see, Reddy, 1979). In the 19th century, this set of metaphors froze as a code view that gave us, first, Morse and telegraphy and, later, computers and the internet. Given the influence of technology, these metaphors were taken up by Saussure and subsequently dominated twentieth century linguistics. However, following Harris (1981), Linell (2005), Love (2004), Kravchenko (2007) and Bickhard (2007) a growing number reject *encodingism*. Far from being a conduit of ideas that are encoded/decoded by minds or brains, language is an ecological whole that functions in many time-scales. It is metabolic or dynamical and, at once, symbolic (Raczaszek-Leonardi 2009). Though part of concerted activity by at least one person, its products are, at once, developmental, historical and evolutionary. Wordings are enacted and, yet, use traditions that are constitutive of the social world. Computers—not living beings—rely on symbolic processes function to encode/decode physical states. Thinking is action: on a machine analogy, total language trains up networks of hypersocial cultural agents that, in their lifetimes, attempt to 'run' languages.

On the conduit view, so-called language 'use' is said to result from the workings of language-systems (e.g., isiZulu, English). Its basis is ascribed to individual knowledge that is represented by a mind or brain. As in Western grammatical tradition, language is described—not around observables (i.e., articulatory activity and pulses of sound)—but by relations between phenomenological forms. Language is thus identified with words, grammars, discourse or constructs that, in some mysterious way, 'reflect' inner thought. Like an artificial system, a brain maps forms onto meaning and, conversely, meanings onto form. Among the problems with any such view is the mereological error of supposing that 'forms' serve brains as input or output. Rather, *people* make and track phonetic gestures that shape how they perceive speech. However, there are no determinate forms *in* the speech wave and we rely on how things are said. As an avalanche of evidence shows, brains exploit rich phonetic information (for review, see Port 2010). Further, we find it hard to track the precise sense of *the words that are actually spoken*. Since connotations affect construal, why does the code view of language persist? Leaving aside the sociology of science, there are two reasons. First, we rely on the language myth (Harris 1981): in everyday life and many institutions language is conceptualised in terms of forms that 'contain' messages. Second, inscriptions use writing systems that invite us to think that form is its essence: since language can be reformatted, we view writing as 'like' language—its essence lies in potential reformatting. Many fall prey to what Linell (2005) calls *written language bias*. It is forgotten that, in themselves, inscriptions lack meaning. Like meanings, forms are *abstracta*.

Later, we adduce further reasons for rejecting code views. However, one of the most compelling is that these reify the phenomenological. True to orthodox science, we prefer to begin with measurable phenomena. By starting with speech coordination we commit ourselves to addressing the goal of how this comes to be *perceived* in terms of forms and meanings. Rather than posit a mental or neural locus, this depends on how language spreads across populations. Perception of form and meaning is a multi-agent phenomenon and, thus, language is distributed (see, Cowley 2011a). It is therefore important to distinguish *languaging* from its products (vocal and other gestures) and the related phenomenology (wordings). Languaging is full-bodied, dialogical, situated and amenable to measurement. Provisionally, it can be defined as "*face-to-face routine activity in which wordings play a part*". Rather than treating forms or meanings as primary, emphasis falls on what we perceive as wordings.[2] Emphasis on coordination allows due weight to be given to the fact that languaging predates literacy by tens-of-thousands or years. By hypothesis, all linguistic skills derive from face-to-face activity or languaging. However, it is only over time that children come to make use of these phenomenologically salient and repeatable aspects second-order cultural constructs (Love 2004). Given the many ways in which they contribute to languaging, they have meaning potential that gives language with a verbal aspect. As Thibault (2011) points out, linguists typically confuse language with second-order constructs. Importantly, in making wordings second-order, we contrast their ontology with that of languaging. Pursuing the contrast, we can use an analogy. Gibson (1979) compared perceiving the world with perceiving pictures. On a card showing Rorschach dots we may see an arrangement of markings and, for example, a dancing bear. Using *discrepant awareness*, we pick up both invariants *of* the picture (e.g., the dots) and invariants *in* the picture (the 'bear'). In languaging too, we pick up invariants of the activity (e.g., how people speak, gesture and use their faces) as well as invariants *in* the activity (e.g., wordings and meanings). Like learning to see pictures, learning to talk draws on discrepant awareness. Just as we see 'things' in pictures, we hear 'things' in utterances. In Cowley's (2011b) terms, we *take a language stance*. On the analogy, this is like learning to see things in pictures while, at the same time, using the body to make one's own (verbal) images. And that, of course, presupposes human agency. Next, we turn to how infants exploit languaging—activity in which wordings play a part—to self-organise their bodies and become human agents who perceive—and make up—wordings.

---

[2] Adults sing, converse, read books, discover new media, and are fooled by advertisements: wordings appear in dreams and silent thoughts. While not languaging, this is also what Love (2004) calls *first-order language*. In all of these activities formal patterns can be said to constrain expressive dynamics.

Human Symbol Grounding

To become fully human, children have to discover how to behave and, among other things, learn how wordings contribute to collective practices. As they become able to play various roles, they benefit from acting and speaking in particular ways. Initially, learning to talk depends on managing concerted action—interactivity—just as much as on wordings. Unlike symbol processors, we use circumstances in co-ordinating in ways likely to achieve strategic ends. At times we act as others expect and, thus, make what count as valid judgements. Practical skills and shared knowledge shape social action. This, of course, connects ontogenesis, training and education. Next, therefore, we sketch how infants use human symbol grounding to sensitise to wordings. Whereas babies rely on interactivity, by 3–4 years of age, second-order constructs (and wordings) exert tight constraints on how children act, think and feel. In tracing how the symbols of a community become part of a person, we depend on what Cowley (2007a) terms *human symbol grounding*.

Though the symbols to be grounded consist in more than wordings, these are foundational.[3] Because they are jointly enacted, they link statistical learning, norms and first-person phenomenology. This triple process begins as a baby's brains sensitise to cultural routines. In its first stage, human symbols are grounded into neural networks. Later, infants learn to act in appropriate ways as symbols-for-a-child are grounded into culture. This second stage is further discussed below. Third, once a child's expressive powers develop, she will start to hear wordings: given brains and culture, symbols will be grounded in first-person phenomenology.[4] Once wordings shape perception, they serve talk about language or, generally, speaking and acting *deliberately*. Over time, the results drive the functional reorganization of feelings, thoughts and actions. Being able to use wordings deliberately is crucial to rational action. As affect-based co-ordination is supplemented by *the said*, children master new routines (and games). Later, children use special modes of action to structure thoughts (Melser 2004). Much is learned by exploiting context to act epistemically (Cowley and MacDorman 2006). Interactivity enables bodies to use real time adjustments to discover 'organisational' constraints. Though neural predispositions influence ontogenesis, they function through concerted activity. Together, infants and caregivers orchestrate by sensitising to affect marked contingencies. They use coaction or, by definition, how one party used the context of another person's action to come up with something that

---

[3] Here a symbol can be defined as a cultural constraint that serves in taking the measure or others and/or in controlling one's behaviour: many symbols are prosodic, gestural or enact what Goffman (1959) calls the interaction order.

[4] There is no clear evidence of when this occurs: however, there is abundant evidence that it is based on the skill of making and tracking phonetic gestures (Fowler and Rosenblum 1991; Fowler 2010). Further, since it is necessary to pretending it is likely that children begin to have the necessary experience in the second half of the second year.

could not have been done alone.[5] At 14 weeks a mother may be able to use, not touch, but the changing context of her body making her baby fall silent (Cowley et al. 2004). The baby attends to repeated action or, in Maturana's terms (1978), how each orients to the orienting of the other. As a result, circumstances are co-opted in strategic (joint) action. Learning to speak is initially separate from co-action. However, caregivers and infants use the rewards of interactivity to share use of contingencies. As infants manage adult displays, they draw on affect or, in Stuart's (Stuart 2012) terms, how *enkinaesthesia* prompts us to orient to the felt co-presence of others. Later, these skills become enmeshed with those of vocalising.

In the first stage of human symbol grounding, brains rely on statistical learning. Before birth brains sensitise to rhythms of voices and languages. Infants show *preferential* response to the mother's voice (and face) and the *right kind of rhythm* (De Casper and Fifer 1980) and, remarkably, a story heard in the womb (De Casper and Spence 1986). While many animals discriminate, babies have skills in expressive co-ordination. Given rhythmic sensitivity, co-action soon falls under the baby's influence. This was first recognised in Bateson's (1971) work on the protoconversations that reveal 'intersubjectivity' (Trevarthen 1979). Context sensitive co-action is also stressed by Bråten (2007). More recently, the ability to co-ordinate expressive movements (including vocalisation) has been traced to grey matter in the brainstem which, before birth, controls morphogenesis (Trevarthen and Aitken 2001). As motivation develops, contingencies prompt a baby to use the rewards to interactivity to anticipate events. By three months, infants gain skills in controlling vocal, facial and manual expression. Norms already play a part in controlling their enkinaesthetic powers. Language and gesture (not to mention music and dance), thus share a neural basis (Willems and Hagoort 2007). As social actors, we rely on controlling expression in time: an infant uses affect to lock on to the movements by others and, by so doing, engages in dance-like co-action. Even congenitally blind infants move their hands to rhythmic patterns (Tønsberg and Hauge 1996). Those who hear normally, however, use its musical properties to discover the rewards of vocalising.

By three months events begin to show the influence of cultural symbols. Infants sensitise to signs of culture: using coordinated action human symbols are *grounded into culture*. Caregivers use infant gaze, smiles and other expressive gestures as indices of local norms that contribute to co-actional routines. Using both biological tricks and adult displays, infants gain a say in events. While infants and caregivers have *fun* together (see e.g., Stern 1971), affect allows interactivity to build contingencies into dyadic routines or *formats* (Bruner 1983). These help infants with when to initiate, what to expect and, of course, when to inhibit. Surprisingly, a three month old may 'do what its mother wants' by falling silent on command (Cowley et al. 2004); in an isiZulu speaking setting, the baby shows *ukuhlonipha* ('respect').

---

[5] Wegner and Sparrow (2007) use experimental work to trace the social and bodily complexity of co-action—and its deep links with our *sense of agency*.

Dance-like interactivity helps the infant re-enact cultural values. This infant changes parental behaviour in ways that induce learning about situated events. Showing ‘respect’ (as we describe it), evokes a feeling tone. Without hearing words (or manipulating symbols), the baby comes to value *ukuhlonipha*. Given co-action, cultural contingencies connect with adult display. In this aspect of human symbol grounding, infant motivations exploit adult *experience*. Even if early normative behaviour uses biology (and neural systems that enable adults to shape infant expression) this is a developmental milestone. Before babies learn to reach for objects, caregivers will sometimes act as if their infants ‘understand what is said’.

As symbols are grounded into culture, a 3–4 month is increasingly adjudged in terms of how well (or badly) she behaves. Given contingencies and rewards, she sensitises to *circumstances*. Instead of needing stimuli or cues, co-action changes how activity is modulated. For many reasons, the focus of development then switches to learning about objects. Late in the first year, however, the child discovers *secondary intersubjectivity* (Trevarthen and Hubley 1978) during Tomasello’s (1999) *9 month revolution*. Bringing social and manipulative skills together, mediated or triadic behaviour emerges. Since language is co-ordinated activity, there is no need for the identification or recognition of inner intentions. Rather, it is sufficient that adults respond *as if* actions were representational. Infants use contingencies (and compressed information) by acting in ways that *seem* intentional. For example, Cowley (2007b) describes a mother who gets a 9 month old to fetch a block. Far from using inferences, the baby co-ordinates with maternal actions that include shifts of gaze, vocalising ‘fetch’ three times and using her whole body as a pointer. Fetching thus appears (and, perhaps, feels) intentional thanks to how the mother’s vocalisation (‘fetch’) encourages norm-based activity. Further, if the child can mimic the sound (fetch), this opens up what Tomasello (1999) calls role reversal imitation. This is facilitated by independent concerns that include infant pleasure in self-produced vocalisations. As babbling shapes articulation, by 12 months, a baby ‘repeats’ syllables. Intrinsic motivations unite with skills and anticipated reactions. In the second year, a baby grasps ‘facts’ linking the normative with the physical. At times, she may do what is wanted. Once a toddler, she follows commands or, strikingly, directs adult attention and actions. She grasps and utters (what adults hear as) ‘words’. As wordings fuse with first-order activity, human agency emerges.

As an infant begins to walk, she is becoming to adopt social roles. While far from reasoning, she draws on virtual patterns and social norms. Given a simple toy, a 12–18 month old will enact cultural expectations. In an unpublished study, a French and an Icelandic child-mother dyad played the ‘same’ game together over several weeks (Sigtryggsdóttir 2007). Whereas the French dyad used this in having fun, the Icelandic partners treated it as goal directed activity. Each baby learned how to elicit rewards. Strikingly, when the Icelandic mother failed to participate, the baby would sometimes self-applaud. Plainly, she exploited—not just affect—but (shared) goals. She has become attuned to the values of her world. Co-ordination enables both parties to use maternal displays of cultural values to organise activity. By participating in routines based in local ways of life, a child learns about fun as well as rationality. Far from relying on sound-patterns, the baby uses rewards that co-vary with what is intended (e.g., 'show respect!'). Quite unknowingly, the child orients to other-directed functions of caregiver verbal patterns. Yet, no 1 year old *hears* wordings. It is only later that utterances come to be heard as utterances *of* patterns. Early on, first-person experience arises in dynamics of co-action (Cowley 2011a, b). Time passes before a child discovers the potential advantages of using wordings as if they were things.

Coming to *hear* verbal patterns changes human (inter) action and perception. While its neural basis lies in using linguistic variability while tracking phonetic gestures, adults use perceptual images of 'words'. All of us have some experience of what is called 'private' thinking by means of public symbols. The skill appears in Piaget's *symbolic stage* when, not by coincidence, infants discover pretending. By hearing more than sound, they discover a 'magical' aspect of language. A 2 year may say (to a banana), "*hello, who's there?*" Without being able to *hear* telephone talk (a remarkable cognitive skill), such pretending could not arise. Given this perceptual learning, a child learns both to get others to do what *she* wants and to use self-directed speech to shape her action. Whereas language sensitive bonobos can follow novel instructions like 'go and put the orange in the swimming pool' (Savage-Rumbaugh et al. 1998), children excel in different skills. Given their biases, they use wordings as social resources. Thus, unlike bonobos, humans share attention for its own sake. By age 3 a human child will not only follow instructions but will language different with peers and in pre-school. She will choose when not to conform: for a child wordings are social affordances. This is just as important in the life of a social actor.

Just as children are not born talking, they are not born rational. Rather, the skills that shape language and reason arise as we identify with aspects of what we hear ourselves say. Co-action prompts infants to orient to local modes of expression. Given a developmental history, layers of agency accumulate together with a history of decision-making. As we redeploy neural resources, we draw on biologically guided interactivity. We learn from a history of anticipating how others will use norms. To act as sophisticated agents training must hone our biological capacities. Human symbol grounding makes us into norm-recognising agents not unlike symbol processors. This depends on individually variable skills in using the language stance to manipulate wordings. Indeed, without this combination, communities would be unable to identify themselves as speakers of a specific language. Without being able to describe words (and rules) as entities that pertain to an autonomous system (e.g. English): we would not believe in abstractions like *minds, selves or societies*.

Using the Resources of Reason

Public language permits *objectively valid* judgements. For Craik (1943), this is exemplified by when, in bridge-building, language is put to symbolic *use*. How this is conceptualised is, of course, a theoretical matter. Since Craik ascribed this to the brain, he viewed language itself as representational. However, the distributed view offers a parsimonious alternative. People learn to refer: they depend on connecting talk about language with languaging (i.e., activity). We learn to take the perspective of the other while linking articulatory patterns with items of shared attention. Once we begin to take a language stance, we hear wordings as wordings that serve to pick out things as things. With practice, we learn to refer or, in other terms, how languaging can be used to pick out objects, persons, events etc.

The skills of a rational human agent depend on both real-time coordination and using the language stance to exploit cultural resources. Given the phenomenological status of wordings, they can be used both literally and in fun. This is because, since they arise from interactivity, they are integrated with action and expression as we enact relationships. In contrast to the fixity of computational symbols, wordings gain effectiveness from flexibility and vagueness. Sense-making arises as they are jointly crafted as persons orient to circumstances and each other. Unlike symbols used in computers, they bear on what people are doing. As part of language flow, interactivity or, in lay terms, *how* we speak is meaningful. Rightly, therefore, many contrast language with man-made codes. Unlike Morse, for example, language is neither monological, disembodied nor dependent on design. Unlike man made codes, language is dialogical (Linell 2005, 2009). Further, given its embodiment, brains ensure that language is enmeshed with action (e.g., Willems and Hagoort 2007). As one thread in coordination, its literal or denotational meaning is often marginal (Maturana 1978). Circumstances add to how, as living bodies coordinate, we make and construe linguistic signs (Love 2004). Dynamics make language irreducible to words, utterance-types, usage-patterns and so on (Harris 1981, Love 2004). As argued here, people—not language—exploit acts of reference: representationalism is not necessary to cognition (e.g., Anderson 2003; Thompson 2007; Stewart et al. 2010). Indeed, computing systems face the *symbol grounding* problem (Harnard 1990): computations are meaningless to machines. Worse still, where grounded by design, symbols fail to pick out facts. No (currently imaginable) robot could ‘know’ that, ‘Put out the light’ is irrelevant to, say, what is in the fridge or a US president’s concerns (see, MacDorman 2007). They face *the frame problem* and, for this reason, robots are increasingly designed in ways that enable them to use people to connect with the world.[6]

---

[6] One of the most remarkable facts about robots is that, already, they use human consciousness: this is exemplified, for example, when they learn to discriminate what we—but not they—see as colours; accordingly, it can be argued that robots are of importance as linguistic informants (Cowley 2008).

In contrast to symbol processing, public and collective behaviour enacts skilled co-ordination. Even infants use actions and utterances as representations (e.g., in pointing). During co-action, adults treat infant doings as intentional. Where the infant identifies relevant features, repetition shapes conventional behaviour. For the adult, the infant acts representationally. To extend its cognitive powers, therefore, the baby tracks contingencies. Indeed, as the cultural world increasingly aligns to joint behaviour, the baby learns about co-action. Later, infants come to *hear* words by using interactivity to track contextual indices of local norms. Further, this is a likely basis for using wordings representationally. By using a language stance they become thing-like entities that sustain *belief* in virtual constructs (Cowley 2011a, b). Accordingly, they can be used both as expected and in transgressive ways. Indeed, both approaches lead to sense-making because of how the results are integrated with coordination. Once we perceive second-order constructs as thing-like, we can generate *thoughts* by modulating how we speak or use a pencil to make pictures. Although partly verbal (and symbolic), interactivity connects bodies, activity and cultural experience. Unlike Morse (or computer programs), language is dialogical, multi-modal and realises values (Hodges 2007). Eventually, wording-based reality links with the resources of an interpretative hermeneutic community.

The spread of language prompts social actors to reproduce society. Within a cultural space, children use interactivity to grasp how symbols serve social action. By coming to anticipate how others speak, they discover Wittgenstein's *agreements in judgement* (1958: § 242). Using coordination, they develop social strategies that depend on connecting words, circumstances, and the music of expression. In this way, unmeasurable virtual patterns take on cognitive power. Like other living things, we depend on compressed (Shannon) information. In Dennett's (1991) terms, humans use *real-patterns* that include not only cases like gravity and colours but also wordings. Because of phenomenological status, we can use for example, as ensembles of norms that reflect on other people's expectations. Davidson's (1997) view of the role of thought and language is similar. He proposes a uniquely human *framework* (p. 27): "The primitive triangle, constituted by two (and typically more than two) creatures reacting in concert to features of the world and to each other's reactions, […] provides the framework in which thought and language can evolve. Neither thought nor language, according to this account, can come first, for each requires the other."

Others concur that thought and language co-emerge from interactions. For Maturana (1978), an agent's sense of self fuses with verbal patterns: *structural coupling* allows new-borns to engage with caregivers. Their languaging soon becomes oriented to *types* of circumstance (and thus a *consensual domain*). This generates (observer dependent) opportunities for sense-making. Gradually, however, perceiving, feeling and acting are integrated with normative aspects of language. As neural functions change, individuals become *speakers*. In our terms, experience allows discrepant awareness to shape skills based on taking the language stance. A child's sense of self uses coordinated action to link cultural resources with individual skills. For example, we talk about talk, develop narratives, and make up autobiographical memory. By using discrepant awareness, we link circumstances with the past, the possible and the future. It is not the brain but, rather, languaging that underpins reference. Even bridge-building integrates symbolic, practical and skill-based knowledge based on a life history of games that make us (more) rational. Using standardisation, dictionary writing and education, (increasing) weight falls on literal meaning. As this becomes familiar, the language stance favours a detached 'point of view' and more body-centred control of thoughts, feelings and actions—provided that we reproduce social 'reality'.

Human Agency Naturalised

In naturalising human agency, we claim that experience of co-ordinating shapes our cognitive, social, and linguistic skills. Thereby we reformulate Durkheim's old claim that the social explains the social, namely by explaining how biological members of sapiens develop the dimensions of sociological agency. While bodies are pre-adapted for cultural learning, interactivity prompts brains to compress information by orienting to verbal patterns, artifacts and norms. We refer by calling up the past, the possible and the future. This is, of course, dependent on institutions and artifacts. Social relations thus underpin reasoning and, of course, skills in making what count as objectively valid or wise judgements. Though we use the results to model social phenomena, their basis lies beyond the brain. We depend on coordinating spontaneously while making judicious use of the language stance and the resources of reason. Though languaging retains its importance to face-to-face thinking, in many other settings, weight falls on treating wordings as wordings. As Piaget (1962) shows, we come to grasp games of marbles or, later, take part in literacy practices. We increasingly use the language stance to participate in the assemblages that enact joint projects. Human agency is partly eusocial. Its develops from a kind of *fission*: as biological infants become persons, a chain of interactivity transforms what they can do. As this happens, they increasingly discern uses for cultural resources that serve in both individual and collective endeavours.

The transformatory power became especially clear when a bonobo chimpanzee, Kanzi, was raised in a human-like environment (see, Savage-Rumbaugh et al. 1998). Not only did he gain from computer access to verbal resources but these were coupled by close attention from human carers. Bonobo symbol grounding made Kanzi strikingly (partly) human-like.[7] The case contrasts with Davis's (1949) description of Anna, whose first years of life lacked social embedding and emotional care. She did not speak, could not eat on her own and never laughed or cried. Lack of human company deprived her of opportunities for learning from how

---

[7] This depends on the observer's point of view: in many respects, Kanzi remains distinctly a bonobo.

people use co-action in orienting to social norms. She never used interactivity in feeling out a cultural world and, as a result, failed to develop the cognitive powers used in social life. Unlike a normal human actor (or Kanzi), her actions were loosely constrained by culture. In short, sociological agency arises as language becomes a dimension of the person. Eighteenth century tradition wrongly plucked words from the world. Language is no transparent medium because, contra Pinker, wordings are not located in the mind (or brain). Rather, they are part of public activity between people, activity that allows even a 14 week old to use co-ordination to show *ukuhlonipha* ('respect'). The baby does not 'encode' meanings or propositions but, rather, learns from the routines of everyday co-action.

A Sketch of Social Fission

For the social sciences, interactivity and languaging are conceptually important. Although everyday language may be the necessary basis for modelling macro-social phenomena, it seems inappropriate to the micro-social domain. The models of social actor theory (Boudon 1981; Coleman 1990; Hedström 2005), like those of code linguistics and the computational theory of mind, ignore the world of embodied, conscious beings. In appealing to social fission, we thus naturalize how the social grounds the social. Rather than treat genes and brains as the origins of reason, we argue that children use interactivity to develop locally appropriate kinds of agency. They draw on experience and, crucially, use the language stance to grasp how people, circumstances and situations vary. Brains and genes predispose infants for cultural learning that, by hypothesis, depends on compressing (Shannon) information. They learn about social (and other) affordances as coordination produces experience with norms, artifacts and wordings. Indeed, the flaws in individual rationality speak strongly *against* ontological individualism. Rationality derives from social relations: it is a feature of the cultural and institutional environment that drives biological humans to make imperfect use of (what count as) objectively valid judgements.

Since symbolic models capture macro-social patterns, biological humans *discover* the resources of reason. Our agency is made and not born; it emerges from both the physical world and affordances such as languages, artifacts and social institutions. Far from centring on a body (or brain), it depends heavily on how languaging enacts social relations. Though, often, we cannot be literal, judicious use of the language stance brings rich rewards. Combined with appraisal and interactivity, we unearth the value of cultural resources. While sometimes acting individually, joint projects tend to dominate our lives: the artificial matters greatly to human agency. It is thus to be expected that coordination serves to make strategic plans. Following Darwinian logic, it is not at all surprising that social affordances are selected as a result of enacting social relations. This may be why most cultures develop, for example, ways of displaying and recognising kinds of *trust* and *reciprocity*. Interactivity gives rise to a selection history that links up languages, institutions and social norms. Human agents develop intuitive or expert skills alongside those based on the resources of reason. Since human *nature* is so flexible, it is an error to use ‘Hobbes Problem’ as evidence for the difficulty of coordination. That said, our limited rationality does create practical problems of aggregation (Spiekermann 2013; Ben-Naim et al. this volume). Rather than view this as a symptom of inherent selfishness, it shows that humans need complex resources that provide results as they move in and out of social aggregates. These make human cognition partly eusocial—much depends on collective modes of action that link the artificial domains of languages, artifacts and institutions.

Interactivity in Human Agency

By acknowledging that cognition cannot be explained by processes within the brain, we move towards a new sociocognitive science. Human agency is constantly re-enacted as interactivity links us with the world. As we do so, we move in and out of social aggregates that draw on languages, artifacts and social institutions. We find our way through the wilds, talk and, for that matter, use computers and develop skills in flying planes. Human agency is not to be identified with the agent. Since it derives from a history of engagement with the world, agency can be traced to four sets of constraints. First, as physico-chemical systems, we exert (and suffer) physico-chemical effects. Second, as living beings, the boundary conditions of our lineage shape the parameters that result from growth, action, learning and development. Third, as human agents, we develop biophysical skills that exploit artificial constraints associated, above all, with artifacts, languages and institutions. Rather than function as boundary conditions, these flexible resources allow us to pursue individual and collective endeavours. Finally, as living subjects, we make and construe artificial affordances. Thus, we are not social actors, languages are not codes and minds are not symbol processors. In rejecting all such organism-centred views, the distributed perspective holds out the prospect of reintegrating biosemiotics, cognitive psychology, linguistics and the social sciences. The core idea is that our becoming can be traced to interactivity that links agents in larger aggregates within a common world. Although creativity gives rise to artifacts, inscriptions and public performances, its basis lies in how biosocial agents mesh temporal scales while using interactivity. Remarkably, it seems that a single sociocognitive system enables brains, languages and societies to conspire in prompting human bodies to make partial sense of the world. This is crucial to the goals of the field. On the one hand, as noted above, we need to clarify how people come to hear and exploit wordings. On the other, this opens up the much broader question of how phenomenological experience links with organization in other time-scales. In short, what can be described in language must be traced, on the one hand, to the rapid time-scales of interactivity and neural processes and, on the other, the slow scales that allow groups to differentiate in ways that drive cultural selection. It is there that fission prompts individuals to become the persons that we are. Just as slow scales constrain faster ones, the rapid processes of interactivity and languaging engender human agency—agents who ceaselessly re-evoke the past to explore the adjacent possible.

Acknowledgements Martin Neumann's contribution was undertaken within the Project 'Emergence in the Loop' (EmiL: IST-033841) funded by the Future and Emerging Technologies programme of the European Commission, in the framework of the initiative "Simulating Emergent Properties in Complex Systems". Stephen Cowley's draws much on the opportunity to contribute to a programme of scientific workshops on *Agency and languaging* (2010–2012) that were organized by the University of Jyväskylä (and sponsored by the Finnish Ministry of Culture).

References

Anderson, M. L. (2003). Embodied cognition. *Artificial Intelligence, 149,* 91–130.

Bateson, M. C. (1971). The interpersonal context of infant vocalization. *Quarterly Progress Report of the Research Laboratory of Electronics, 100,* 170–176.

Bickhard, M. (2007). Language as an interaction system. *New Ideas in Psychology, 25,* 171–187.

Boudon, R. (1981). *The logic of social action: An introduction to sociological analysis*. London: Routledge.

Bråten, S. (2007). Altercentric infants and adults: On the origins and manifestations of participant perception of others' acts and utterances. In S. Bråten (Ed.), *On being moved: From mirror neurons to empathy* (pp. 111–135). Amsterdam: John Benjamins.

Bruner, J. (1983). *Child's talk: Learning to use language*. New York: WW Norton Press.

Coleman, J. (1990). *Foundations of social theory*. Cambridge: Belknap.

Cowley, S. J. (2007a). How human infants deal with symbol grounding. *Interaction Studies, 8,* 81–104.

Cowley, S. J. (2007b). The cradle of language: Making sense of bodily connections. In D. Moyal-Sharrock (Ed.), *Perspicuous presentations: Essays on Wittgenstein's philosophy of psychology* (pp. 278–298). Houndmills: Palgrave MacMillan.

Cowley, S. J. (2007c). Cognitive dynamics and distributed language. *Language Sciences, 29,* 575–583.

Cowley, S. J. (2008). Robots: The new linguistic informants? *Connection Science, 20,* 359–369.

Cowley, S. J. (2009). Language flow: Opening the subject. *Cognitive Semiotics, 4,* 64–92.

Cowley, S. J. (2011a). *Distributed language*. Amsterdam: John Benjamins.

Cowley, S. J. (2011b). Taking a language stance. *Ecological Psychology*, *23*(3), 185–209.

Cowley, S. J. (in press). Naturalizing language: Human appraisal and technology. *AI & Society.*

Cowley, S. J., Moodley, S., & Fiori-Cowley, A. (2004). Grounding signs of culture: Primary intersubjectivity in social semiosis. *Mind, Culture and Activity, 11,* 109–132.

Cowley, S. J., & MacDorman, K. F. (2006). What baboons, babies and tetris players tell us about interaction: A biosocial view of norm-based social learning. *Connection Science, 18,* 363–378.

Craik, K. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.

Cronin, B. (2004). Bowling alone together: Academic writing as distributed cognition. *Journal of the American Society for Information Science and Technology, 55,* 557–560.

Davies, K. (1949). *Human society*. New York: Macmillan.

Davidson, D. (1997). Seeing through language. *Royal Institute of Philosophy, Supplement, 42,* 15–28.

Dennett, D. C. (1991). Real patterns. *Journal of Philosophy, 88,* 27–51.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mother's voices. *Science, 208,* 1174–1176.

DeCasper, A. J., & Spence, M. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant, Behavior and Development, 9,* 133–150.

Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency". *Phenomenology and the Cognitive Sciences, 4,* 429–452.

Durkheim, E. (1895[1998]). *Rules of sociological method*. Basingston: Macmillan.

Fowler, C. A., & Rosenblum, L. D. (1991). The perception of phonetic gestures. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 33–59). Hillsdale, NJ: Lawrence Erlbaum.

Fowler, C. A. (2010). Embodied, embedded language use. *Ecological Psychology, 22,* 286–303.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Goffman, E. (1959). *The presentation of self in everyday life*. New York: Doubleday Anchor.

Harnard, S. (1990). The symbol grounding problem. *Physica D. Nonlinear phenomena*, *42*(1–3), 335–346.

Harris, R. (1981). *The language myth*. Oxford: Duckworth.

Hedström, P. (2005). *Dissecting the social*. Cambridge: Cambridge University Press.

Hodges, B. (2007). Good prospects: Ecological and social perspectives on conforming, creating and caring in conversation. *Language Sciences, 29,* 584–604.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge MA: MIT Press.

Järvilehto, T. (1998). The theory of the organism-environment system: I. Description of the theory. *Integrative Physiological and Behavioral Science, 33,* 321–334.

Järvilehto, T., Nurkkala, V. M., & Koskela, K. (2009). The role of anticipation in reading. *Pragmatics & Cognition, 17,* 509–526.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18,* 513–549.

Kravchenko, A. (2007). Essential properties of language: Why language is not a digital code. *Language Sciences, 29,* 621–650.

Linell, P. (2005). *The written language bias in linguistics: Its nature, origins and transformations*. London: Routledge.

Linell, P. (2009). *Rethinking language mind and world dialogically: Interactional and contextual theories of human sense-making*. Charlotte NC: Information Age Publishing.

Locke, J. (1690/1975). *An essay concerning human understanding*. In P. H. Niddich (Eds.), Oxford: Oxford University Press.

Love, N. (2004). Cognition and the language myth. *Language Sciences, 26,* 525–544.

MacDorman, K. (2007). Life after the symbol system metaphor. *Interaction Studies, 18,* 143–158.

Maturana, H. R. (1978). Biology of language: The epistemology of reality. In G. A. Miller & E. Lenneberg (Eds.), *Psychology and biology of language and thought: Essays in honor of eric lenneberg* (pp. 27–63). New York: Academic Press.

Melser, D. (2004). *The act of thinking*. Cambridge, MA: MIT Press.

Rączaszek-Leonardi, J. (2009). Symbols as constraints: The structuring role of dynamics and self-organization in natural language. *Pragmatics & Cognition, 17,* 653–676.

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284–297). Cambridge: Cambridge University Press.

Ross, D. (2007). H. sapiens as ecologically special: What does language contribute? *Language Sciences, 16,* 710–731.

Pattee, H. H. (1969). How does a molecule become a message? *Developmental Biology Supplement, 3,* 1–16.

Pattee, H. H. (1997). The physics of symbols and the evolution of semiotic controls". In Proceedings from the Workshop on Control Mechanisms for Complex Systems: Issues of Measurement and Semiotic Analysis. Santa Fe Institute Studies in the Sciences of Complexity. Redwood City, CA: Addison Wesley.

Piaget, J. (1962). *Play, dreams and imitation in childhood*. New York: Norton.

Port, R. (2010). Rich memory and distributed phonology. *Language Sciences, 32,* 43–55.

Raguso, R., & Willis, M. (2002). Synergy between visual and olfactory cues in nectar feeding by naïve hawkmoths Manduca Sexta. *Animal Behavior, 64,* 685–695.

Savage-Rumbaugh, S. E., Shanker, S. G., & Taylor, T. J. (1998). *Apes, language and the human mind*. Oxford: Oxford University Press.

Sigtryggsdóttir, T. (2007). Guiding attention leads to changes in child's manifest practical understanding. Unpublished dissertation, University of Hertfordshire.

Spiekermann, K. (2013). Judgement aggregation and distributed thinking. In S.J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition Beyond the Brain* (pp. 31–51) (First edition). London: Springer.

Steffensen, S. V. (2011). Beyond mind: An extended ecology of languaging. In S. J. Cowley (Ed.), *Distributed language* (pp. 185–210). Amsterdam: John Benjamins.

Stern, D. N. (1971). A micro-analysis of mother-infant interaction: Behaviors regulating social contact between a mother and her three-and-a-half-month-old twins. *Journal of American Academy of Child Psychiatry, 10,* 501–517.

Stuart, S. A. J. (2012). Enkinaesthesia: The essential sensuous background for co-agency. In Z. Radman (Ed.), *The background: Knowing without thinking* (pp. 167–186). Houndmills, Basingstoke: Palgrave Macmillan.

Stewart, J., Gapenne, O., & Di Paolo, E. (2010). *Enaction: Toward a new paradigm for cognitive Science*. Cambridge MA: MIT Press.

Thibault, P. J. (2011). First-order languaging dynamics and second-order language: The distributed Language view. *Ecological Psychology, 23,* 210–245.

Thompson, E. (2007). *Mind and life: Biology, phenomenology and the sciences of mind*. Cambridge MA: Harvard University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge: Cambridge University Press.

Tønsberg, G. H., & Hauge, T. S. (1996). The musical nature of prelinguistic interaction. The temporal structure and organization in co-created interaction with congenital deaf-blinds. *Nordic Journal of Music Therapy, 5,* 63–75.

Trevarthen, C. (1979). Communication and co-operation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech* (pp. 321–347). Cambridge: Cambridge University Press.

Trevarthen, C., & Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (Ed.), *Action, gesture, and symbol* (pp. 183–229). New York: Academic Press.

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory and clinical applications. *Journal of Child Psychology and Psychiatry, 42,* 3–48.

Wenger, D., & Sparrow, B. (2007). The puzzle of coaction. In D. Ross, D. Spurrett, H. Kinkaid, & L. G. Stephens (Eds.), *Distributed cognition and the will* (pp. 17–41). Cambridge MA: MIT Press.

Willems, R., & Hagoort, P. (2007). Neural evidence and the interplay between language, gesture and action: A review. *Brain and Language, 101,* 278–289.

Wittgenstein, L. W. (1958). *Philosophical investigations* (2nd ed.). Blackwell: Oxford.

Chapter 10
Living as Languaging: Distributed Knowledge in Living Beings

Anton Markoš, Jana Švorcová and Josef Lhotský

Abstract We trace life at different levels of organization and/or description: from protein ecosystems in the cell up to the cohabitation of individuals within and between historically established lineages. Ways of such cohabitation depend on experience of particular guilds or aggregates; they cannot be easily foretold from any basic level of description, they are distributed across all levels, and across all members of the community. Such phenomena of interactivity constitute a lived world which, we argue, represents a genuine analogy with domains of human cultures and languages. We draw an analogy with three levels of meaning as defined by Rappaport (2010) and make an attempt to show that life and languaging are virtually analogous. Contributions to this volume show that cognition arises not only 'in the head', but also as the result of living in a network of interactions—in the medium of languaging; language and languages cannot be separated from languaging, and our joint activities make sense because of how we concert our doings in a culture or what Thibault (2011) terms a social meshwork. Outcomes of such doings often depend also on differences that people find as meaningful cues to perform expertly or to construe wordings in a particular way. In other words, much depends on patterns that are extracted by living beings that dwell in a historical world of bodily experience, and of the community into which they are rooted. Indeed, in the context, these ideas will not seem controversial; however, in what follows, we propose taking a further step: we propose that analogical processes help *all* inhabitants of the biosphere/semiosphere to become valuable members such of living networks. Our approach may look as yet another contribution to the long list of holistic theories that compete without success with the reigning reductionist paradigm of biology. We, however, do not deny the explanatory power of

---

A. Markoš (✉) · J. Švorcová · J. Lhotský
Department of Philosophy and History of Science, Faculty of Science,
Charles University Prague, Viničná 7, 128 44 Praha 2, Czech Republic
e-mail: markos@natur.cuni.cz

J. Švorcová
e-mail: jana.svorcova@natur.cuni.cz

J. Lhotský
e-mail: lhotsky@natur.cuni.cz

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_10

contemporary biological theory: by stressing the role of historical bodily experience and of the “cultural” role of communities we strive towards a fuller understanding of life phenomena, much along the line the linguists undertook from the vocabularies through semiotics up to languaging. We invite the reader to take an excursion from the “central dogma” and neodarwinian explanation of evolution, towards what we believe is a more complete view of the living, that extends through 9 orders of magnitudes (or “73 octaves of nature‘s music”, as poetically expressed by Ho (1993)) and from nanoseconds to 4 billions of years. Our extension to the distributed view is to argue that what goes for cognition and language also applies generally to life.

Levels of Meaning

Biologists have much to gain from considering how human cultures exploit what have been termed various ‘levels’ of meaning. Here, we take inspiration and a leading thread in the book by Rappaport (2010), *Ritual and religion in the making of humanity*; we shall exploit its paraphrase “Ritual in the making of species”, still by following Rappaport’s argumentation that was intended for the human race only.

Rappaport invites us to acknowledge human cultures as featuring three levels of meaning. Our paper will take the view that the kinds of systems that we find in molecular biology bear remarkable similarities. (1) L*ow-order meaning* is based in differences that can be found in the everyday semantics: thus *rat* differs from both *mouse* and *rate* (in spelling as in pronunciation). Plainly, science is most comfortable with this kind of meaning, and we shall investigate some features of this level in biology. (2) In the *middle-order* of meaning, a person is able to make and construe “similarities hidden beneath the surfaces of apparently distinctive phenomena” (Rappaport 2010, p. 71). While types may still appear, they are now associated with various kinds of metaphors and signs. This is the level of biosemiotics and biohermeneutics, and we took casual examples how an individual construes its body and its umwelt at this level of meaning. Finally, (3) *high-order meaning* is “grounded in identity or unity, the radical identification or unification of self with other” (p. 71); in dealing with this, we look beyond models that depend on the regular appearance of discrete types and draw on what we think of as “experience of being” and our sense of belonging in a community. Rappaport concludes (*caveat lector*, he speaks about human condition!): “The distinctions of low order meaning, lodged in language, divide the world into discrete objects; the recognition of similarity constituting middle-order meaning makes connections among those objects; high-order meaning unifies the world into wholeness. Middle-order and high-order meaning may thus prevail, at least from time to time, over the experiences of fragmentation and alienation that are, possibly, concomitants of language’s objectifying powers, but it is important to note that *the three levels of meaning do not always live easily together*. Naive scientism and naive rationalism tend to deny the validity of middle- and high-order meaning, and it is ironically interesting that information may be the enemy of meaningfulness. Conversely, untempered commitment to middle- and high-order meaning may ignore crucial distinctions in the natural and social world." (Rappaport 2010, p. 73).

Let us explain the three levels on a Biblical parable: Ezequiel cites the Lord as declaring: "I have no pleasure in the death of the wicked" (33, 11). While the verse may be new to some readers, it has been cited and interpreted in numerous sermons, moral debates and literary contexts. Yet, we suggest, none of hypothetical readers, naïve or learned, is likely to have considered the sentence in terms of the following syllogism:

p: God has no pleasure in the death of the wicked
q: Mrs. A is wicked
$p \rightarrow q$: Mrs. A is immortal

Yet this what the sentence means in plain English! If language functioned like a unidirectional code, it would evoke Rappoport's low-order level of meaning. Why, then, do we not attribute immortality to Mrs. A? Our case is that the networks or paraphernalia of our civilization leads us to read the verse in relation to higher orders of meaning. This is not based on interpretation of the discrete signs at all: we *feel* that God would not grant such things; our cognitive biases link "death" with "damnation". Readers who are familiar with the Babylonian exile will place the prophet's words yet in another context that derives from our understanding of the original, our acquaintance with Middle Eastern realia, and our own cultural milieu. Still nothing of this will explain, why 2600 years after being written, the verses do still appeal to people in, for example, Central Russia or Arizona.

Middle- and high-level layers of meaning are often seen as bound up with hermeneutics, therefore as part of the humanities, as distinguishing humans from the rest of the biosphere. Biology, it is assumed, can be studied independently of history, cultural contexts, language-like patterns, experience and so on. Yet, in the "The chaos of everyday life", Rappaport suggests (2010, xvi), "stability is bound up with the social facts of a shared collective existence". Not only do we depend on history, the reiteration of forms and experience but we also draw on, clichés, metaphors, ritualized activities and even strange assumptions. In Umberto Eco's terms: "…it is impossible to communicate without putting something into the *background frame of mutual agreement and assuming that the other is able to access this presupposed knowledge*. Otherwise, each speech event would require a complete restatement, with the result that there would be no time to say, or listen to, anything. This is clearly too great an extension for a presupposition as a sentence phenomenon, since the utterance of even the simplest sentence can presuppose all the world in this sense."(Eco 1994, pp. 228–229, emphasis added)

In turning to how language and cognition play out 'in the wild', such ways of meaning appear less exotic. Human actions are situated in a normative world where bodies use learning (and other interactional products) to co-ordinate internal and external structure. People, moreover, do this collectively as 'co-acting assemblages' (Cowley & Vallée-Tourangeau 2010, p. 469). Persons-embedded-in-action and/or-interaction resemble to a "field force", built and rebuilt continuously by inhabitants of a "field" that was inherited from those who long since passed away. Heidegger (1982, 1995) calls this Being-with-others (*Mitsein*) in a Country (*Gegend*). This country is moulded by, on the one side, tectonic forces and, on the other, efforts by those who share their being in the here and now. In this way, a countryside or culture is able to evolve across innumerable generations. If this is indeed the basis of cognition, it cannot be traced to simple encodings. This is because, in coming up with thoughts, people draw on distant factors—like the words of an ancient prophet in our example above. Can we, however, generalize from human experience to biosphere, without committing the flaw of anthropomorphization?

In what follows, we show that biological codes such as the DNA script, intracellular and intercellular signal systems and ecological cohabitations also have a strange duality. While participating in unidirectional processes, they also inhabit a 'country' of messages and lineages. The "scientific" treatment of "biological syllogisms" applies in artificial, laboratory settings. Like thinking and sense-making, living processes and their evolution depend on interactivity, a process Kirsh (2010, p. 441) defines as a "back and forth process where a person alters the outside world, the changed world alters the person, and the dynamic continues". Many challenge such a view: in line with the central dogma of molecular biology (see below), many focus on the lowest Rappoportian level. It is hoped that higher levels of meaning are emergent phenomena that can be explained by focusing on such a lowest level of description. It is as if, in studying life, one can ignore the role of living beings. Yet, in Western culture and, thus, the humanities, this view is common; even Rappaport (2010, p. 10), who should know better, concurs: "Non-human systems are organic systems constituted largely by genetically encoded information. Human systems are cultural-organism systems constituted by symbolic (linguistic) as well as genetic information." In challenging this, we aim to rescue the study of life itself from the no man's land that lies between sciences and humanities. Our axiom be: *All living systems are cultural-organism systems constituted by symbolic (linguistic) as well as genetic information."*

The Low Order

Central dogma of molecular biology: Genetic information inscribed sequentially in nucleic acids (DNA or RNA) is decisive for the structure and function of proteins; proteins, however, have no means for feedback, that is, they cannot implement any changes in the genetic information. As proteins represent the principal agency in construction of the body (the phenotype), it follows, that all relevant information how to build a body is inscribed in the form of a linear string of building blocks ("characters") constituting the chain of nucleic acids.

In biology, by adopting the so called "Central dogma", the focus has fallen exclusively on the low order of meaning (see any modern textbook in molecular or cell biology, e.g., Alberts et al. 2007; the reader who is acquainted with basics of molecular biology can safely skip this section). It claims that information flow in biological systems is unidirectional, from script encoded in DNA to proteins to higher levels of organization. Hence, the basic level of description of any living being is its master copy of DNA containing "data" and "programs" how to build the body. Even instructions how to construct the "hardware" (or better, wetware) of the body must be in its entirety encoded in the genetic script (its "wording" is called *genotype*). In the process of *transcription*, parts of DNA information (about 30,000 "genes", constituting about 2–4% of DNA in human cells) are transferred to much shorter strings of RNA; one class of RNAs (messenger or mRNAs) serves for *translation* of information into a string of a particular *protein* (more about proteins see below). The translation rules—the genetic code—extend the realm of chemistry: the code was established in evolution by natural conventions (Barbieri 2008a, b). Several thousands of different kinds of proteins constitute the lowest level of cellular agency (higher levels being multiprotein complexes, membranes, organelles, and other structures) responsible for metabolism, locomotion, cell division, but above all for the extremely reliable *replication*, that is copying of DNA master copy, to distribute it to daughter cells, and, of course, also "to read" it in the transcription and translation processes described above. The assembly of agencies and structures constitute cells, and cells build multicellular bodies; how such an assembly looks like, that is what is its *phenotype*, is *primarily* the function of the genetic script implemented in DNA. To repeat again: there is *no reverse flow of information*—no feedback from the world or flesh into the script (see, e.g., Monod's (1972) the classical treatise *Chance and necessity*). Phenotypes, and other structures of biosphere web, essentially obey, as if verbatim, the genotypic instructions. There are no pterodactyls in contemporary biosphere, because no pterodactyl genotypes operate in contemporary cells, they were lost long ago. Flaws and paradoxes of the theory came to light relatively early (see, e.g., Hofstadter 1979), yet the debates on the topic often end with a mantra "In principle the Central dogma holds". The problem, of course, lies in the fact that all living beings have been born of living beings, they do not start from scratch like crystals, flames, neither are they products of assembly lines. Bodies and they genetic scripts are co-extensive, neither is "primary" or more basic.

The contemporary neo-Darwinian paradigm, however, draws on the Central dogma. Replication of DNA is highly, but not absolutely reliable (typos, and even more serious disruptions may occur due also to external factors), hence, genes in a population may come out in slightly different "spellings" called *alleles* (likewise, "program" and "programme" represent two alleles of the same word). Different alleles (and coalitions of alleles) result in proteins, that is also bodies (phenotypes), slightly differing (in this or that respect) from other individuals present in the population, and such differences may influence the *fitness* of that particular body—in terms of the amount of its descendants. The body is, then, a vehicle to transmit its burden of its alleles into the next generation: the fitter the vehicle, the higher the frequency of particular allele(s) in the population in the next generation. The fitness is determined by *natural selection* in the external environment: Because of the Central dogma, natural selection acts on the carnal vehicles, whereas the gist of evolution is in transferring pure information as inscribed in DNA. For more succulent version of the story see e.g. Dawkins (1976, 1982).

What is important for our further exploration is the fact that the Central dogma and neo-Darwinism models presuppose the concept of a 'basic level' in description of the living. Living beings are viewed as passive machines that are designed to transfer their "software" into their progeny. One way of countering such views is to exploit the language metaphor of life (Markoš and Faltýnek 2011; Markoš et al. 2009; Kleisner and Markoš 2005, 2009). Rather than dwell at the lowest level of meaning, we look beyond models that depend on discrete types and, in so doing, show the relevance of higher levels of meaning to the realm of living. As we argue below, living systems draw also on ecological (or oiko-logical) aspects of meaning. It is our view that recognition of their historical basis is necessary to placing life in a coherent system of knowledge that brings out the continuities that link it with the many human worlds that unfold within a cultural meshwork. As there is no external agency steering the living processes and their evolution, we would like on some examples that life acts as its own designer (Markoš et al. 2009), that is, the lowest level of meaning will not satisfy the task. Yet in other words: the agency driving both ontogeny and evolution is distributed across many levels of bodily organization, with no primary, or central, steering controls.

The Middle Level

One way of countering central dogma with its basic level of description is to exploit the language metaphor of life (Markoš and Faltýnek 2011; Markoš et al. 2009; Kleisner and Markoš 2005, 2009). Rather than dwell at the lowest level of meaning, we look beyond models that depend on discrete types and, in so doing, show the relevance of higher levels of meaning to the realm of living. In so doing, we face opposition from both the sciences and the humanities (see, e.g., Heidegger 1995). However, we see no need for this: accordingly the paper aims to show that, in contrast to views associated with the kind of logic associated with the central dogma, living systems draw on ecological (or oiko-logical) aspects of meaning. It is our view that recognition of their historical basis is necessary to placing life in a coherent system of knowledge that brings out the continuities that link it with the many human worlds that unfold within a cultural meshwork.

We pursue the "language metaphor of life" beyond the affairs of *Homo sapiens*, into communities of living entities. In arguing that it is essential, we show that history and experience matter to intracellular processes, cells living in a body, members of a species and even ecosystems. Life depends on, or better dwells in, cultures or, in Kauffman's (2000) terms, *biospheres* made up of populations of cooperating *autonomous agents*. Many of the properties of languaging (Markoš 2002; Markoš et al. 2009; Markoš and Švorcová 2009; Markoš and Faltýnek 2011) appear in communities or guilds of living entities: the processes that sustain life are *radically distributed in that they depend on 'memory' that is inseparable from their surroundings*. Living beings are not produced or created *ex nihilo* like crystals or tornadoes: they are *born* into already present “biospheric fields”. Parental individuals (and the community) give birth to beings that develop in a pre-existing domain of rules, values, heuristics and ways of doing things. Hence, besides the genetic script and the body that harbors its patterns, we emphasize the third factor—the community. We now focus our approach around four examples: (1) the intracellular “ecology” of the protein world; (2) epigenetics; (3) symbiosis and symbiogenesis; and (4) the new science of evolution-development, affectionately known as evo-devo (Carroll 2005; Gilbert and Epel 2009).

Proteins as Agents at the Molecular Level

In our view it is difficult to understand life without considering properties of the protein community. Proteins are huge molecules. By comparison, if we treat the “size” of a hydrogen atom as 1 and that of water as 18, a protein averages at about 40,000 (10–100,000). Each of their “building blocks”, an amino acid, has a size of around 100. Proteins are always synthesized as linear chains consisting of aperiodic sequences that are constituted by 20 different species of amino acids; the chain is synthesized according to a sequence of particular sections in DNA called genes. Genes are first copied (transcription) into “messenger RNA” which is translated in accordance with a sequence of instructions (the genetic code) into the amino acid chain that constitutes the protein. It should be pointed out that the whole process is catalyzed and steered by pre-existing protein “machinery” that, in its turn, arose also from the transcription-translation process.

The resulting native protein chain shows sensitivity to a particular train of amino acids by wrapping onto itself and creating a 3D molecular protein molecule. Given the view that all necessary information is contained in the DNA (e.g., Monod 1972; Anfinsen 1973) many thought that a one-dimensional codon sequence unequivocally determined both the chain and the shape of the molecule. On this view, since proteins are the “basic building blocks” of the cell, the shape of cells and multicellular beings is to be traced to the code of a genetic script. In fact, a protein molecule can attain an astronomic number of different shapes. In a given case, however, their embedding in a cellular environment will ensure that only a limited (“meaningful”) number are attained (Fig. 10.1). Misfoldings are quickly repaired, or removed—by the cell’s protein assembly apparatus.

Most proteins possess *binding site(s)* for a specific *ligand* (a small molecule or specific shapes recognized on macromolecules like DNA, RNA, sugars, or other proteins). On binding the ligand, the molecule *does* something by *changing its shape* (conformation): it may change the chemical nature of the ligand (enzyme), bind an antigen (antibody), transfer molecules across a membrane (channels, pumps); pass or amplify signals (receptor); etc. These are not coding processes (based on input-output relations) but rather *performances* that change the protein molecule’s shape while binding its ligand(s). Every protein depends on being able to change its shape upon interacting with its environment. A mammalian cell

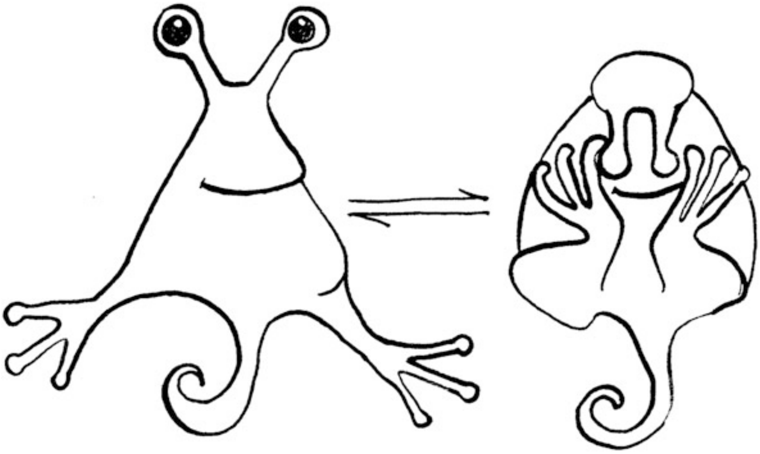


Fig. 10.1 Two possible conformations of a protein molecule

contains about 30,000 genes of which, in a given cell, 10,000 are typically ‘read’. However, the set of actual protein shapes in the cell is much larger: as explained below, this depends on the protein protein ecosystem into which new proteins are born (for more detail and self-explanatory cartoons, see Alberts et al. 2007).

In order to attain proper shape a great many nascent proteins depend on “chaperons” (Rutherford and Lindquist 1998; Bergman and Siegal 2003; Taipale et al. 2010; Fig. 10.2). The set of chaperon proteins thus become major regulatory “hubs” that, in different regimes, regulate the cell’s crowded protein network by means of fine-tuning (Taipale et al. 2010). In a broader context, not only chaperons but all pre-existing structures and protein assemblages can play formative roles in the environment where a protein molecule is born (e.g., Good et al. 2011). Hence the decision of the context in which the protein is to work is by no means local; it results from the ecosystem of cell “inhabitants”. Thus, without any need for central control, proteins function as a distributed meshwork of complex system.

Shape transitions are necessary to protein function. To perform a specific action each must take on a conformation that gives exquisite sensitivity in distinguishing and binding its ligand. On binding, the conformation changes and, by so doing, sets off special operations on or with the ligand. It may, for example, be chemically transformed or transported; a change in conformation may switch a signalling pathway or, perhaps, set off protein-protein binding. The changing conformation can prompt a functional complex to perform a task. The effects of such a change are sketched in Fig. 10.3. During such functions, the protein’s performance (speed, efficiency, etc.) may be fine-tuned by the protein protein ecosystem. While about one tenth of proteins in the cell are bound to “housekeeping” functions (e.g., respiration, food intake, or special syntheses), the others act as a regulatory, information processing network that make subtle responses to whatever happens to the cell.

The function of a protein is *distributed* in that it does not rely on predetermined features alone; it also draws on historical (evolution, ontogeny, given cellular

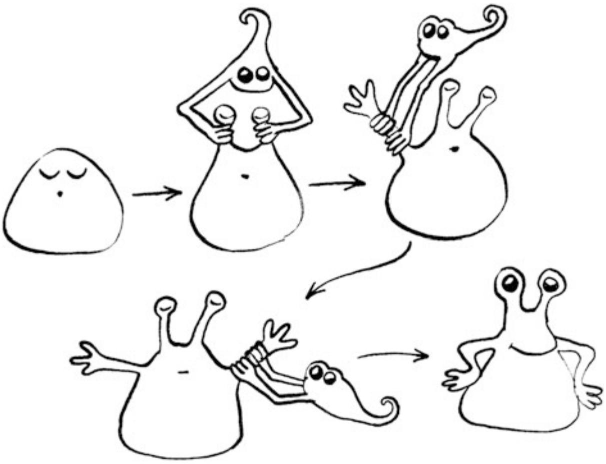


Fig. 10.2 The action of a chaperone on the nascent protein (in many cases, contact with a chaperone is required across the whole lifetime of a given protein)

context) and ad hoc contingencies (e.g., temperature), or, in short, on the *experience* of the cell and organism. Such a statement somewhat complicates the straightforward model of evolution described in Fig. 10.4.

Undoubtedly, evolution draws on random change mutation in the genetic script. As described in every textbook, this leads to alterations in the sequence of protein-coding or regulatory sections of DNA. As a result of change in respective DNA sections, a protein may alter its performance; mutations in the regulatory sequences may also place proteins in new contexts by, for example, altering the timing of ontogenetic gene expression. Changes in the setup of protein network (ecosystem) can have far-reaching consequences for a cell, an individual's appearance (phenotype) and, indeed, for the ecosystem in which it lives. There is, moreover, a second kind of evolutionary change. A whole network of proteins may be induced to change its Performance by external factors such as temperature, nutrition, epidemic, that change the appearance and performance of its bearer. If the whole population is the target of such a change, an unaltered genetic script may nonetheless present a new "attitude towards the world" (see, Matsuda 1987; Hall et al. 2004). Given these two modes of evolution, the network has distributed functions. This is important because, contra the central dogma of biology, this can*not* be traced to inscriptions in genetic code, indeed, it depends on non-local factors that are co-dependent with biochemistry, molecular configurations, function and evolutionary effects. If epigenetic causation (often reversible from the beginning) takes many generations it may even come to be fixed by genetic algorithms (e.g. Waddington 1975, Rutherford and Lindquist 1998). Next, therefore, we turn to how cells develop.

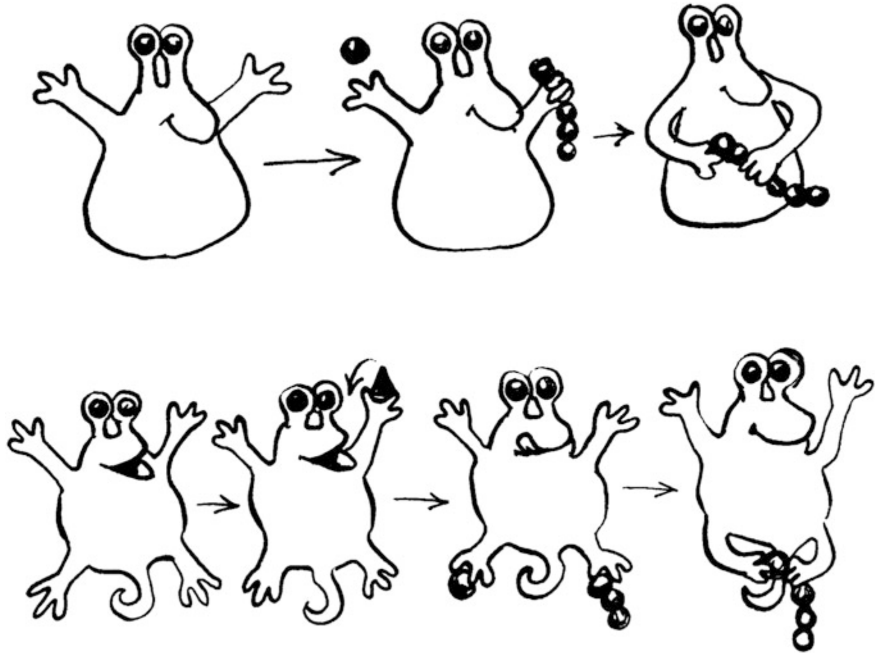


Fig. 10.3 In the *top row* a given protein functions by adding a molecular element to a growing chain. The protein has binding sites to both ligands (the monomer and the chain). Thus, when ligands bind onto specific sites, they induce unifying changes in conformation. In the *lower row* a protein molecule couples with an energy source that enables the inactive conformation to attain the receptive shape required for work (if ligands are available and bound to appropriate sites)

Epigenetics in the Lives of a Cell

Now, we shift our focus from distributed control to consider how a cellular system attunes its current needs by using the 'wording' of genetic texts. We find a sophisticated process that is reminiscent of the subtle use of alternations that "accent" an alphabet's basic letters (e.g. 'a') by marking them as (for example) á, ä, à, â, ã, ǻ, å, ặ, ã̄, etc. While from the point of view of the original Latin such modifications look bizarre, they perform many functions. Even if the differences do not matter at one level (e.g., in e-mails), there are substantial differences at others (e.g., in German, Bar/Bär are different words as are tacher/tâcher in French). In the cell, marks are (reversibly) placed onto DNA or proteins and thus altering the "text" that influences how proteins perform.

Epigenetic use of a diacritic like process is far from simple. They ensure, for example, that cells which inherit the same basic 'text' from the zygote can develop into, say, a liver or a brain. As different sets of proteins contribute to the relevant epigenetic processes, organ formation depends on the highlighting and suppression of different parts of genome and/or proteome. There are two key processes in the

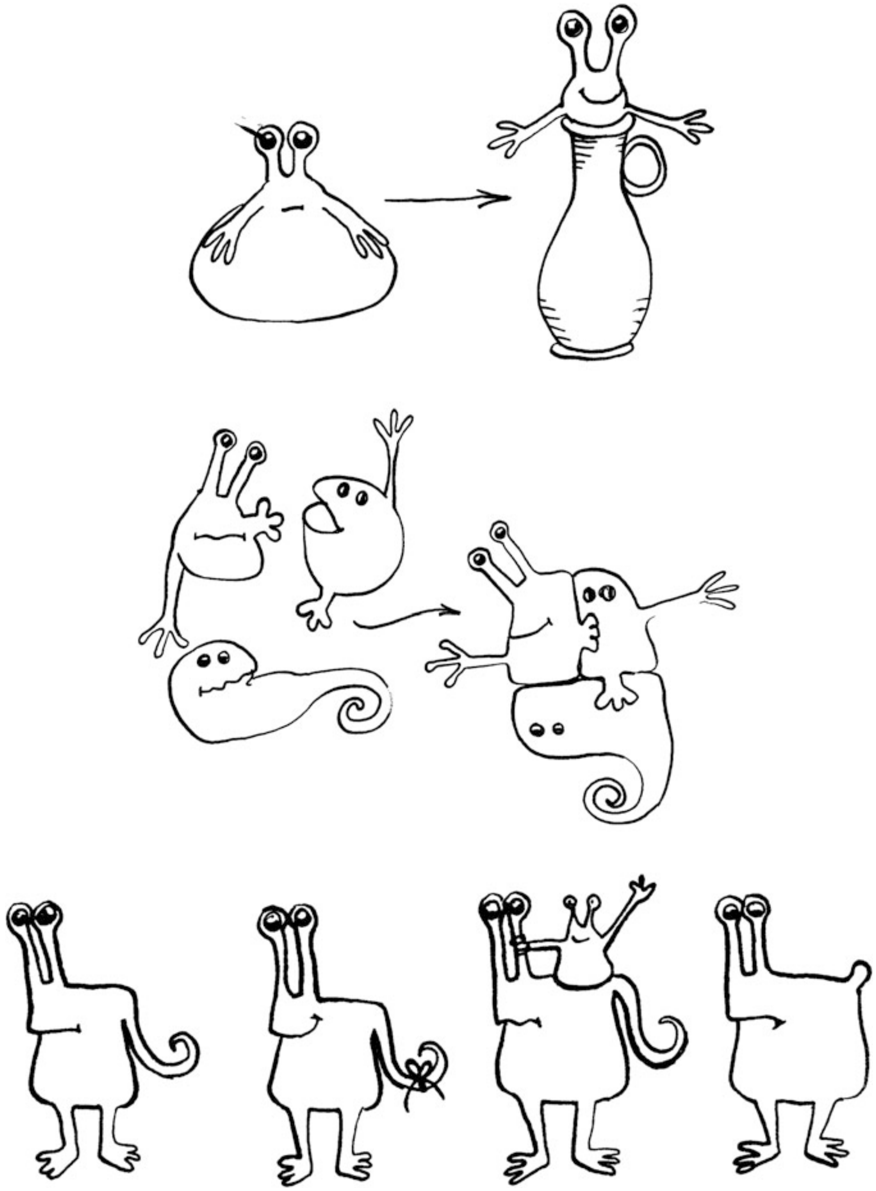


Fig. 10.4 The performing conformation can be also attained by embedding protein into a structural and/or functional context of a specific environment, or can be delicately (or less delicately) and reversibly modified by specific action from its environment

cell nucleus that help cells (and cell lineages) to remember their spatial and temporal coordinates.[1] The first of these adds chemical marks (i.e., “diacritics”) to DNA molecules. These alter how the script is treated, read and/or understood. The second process uses the organization of the cell’s nucleus or scaffolding structures known as nucleosomes. Both processes are tightly interwoven, and deeply influence the cell’s orientation and workings.

DNA Markings and Genomic Imprinting

One-dimensional molecules of DNA are often compared to a letter string written in 4 “characters” A, C, G, and T (chemically—nucleotides). On this linear model, the chemical modification of characters resembles human use of diacritics. The commonest of these modifications (methylation) applies to the C character or cytosine in the DNA string. For some mechanisms, nothing has changed (e.g., DNA replication uses the 4 bases); however, for others, the string features a fifth character in the string. Such modifications are reversible in the sense that another battery of enzymes can remove the “diacritics”. The method allows methylation to influence the accessibility—and transmissibility—of specific DNA strings. In a reading metaphor, it enhances or hides the text from the performing proteins (see Fig. 10.5). The reversible process of DNA modification can profoundly influence a cell’s internal milieu. This is because it is only by binding proteins to regions of a DNA string that the encoded ’message ’ can be transmitted to the body-world. Thus, if the functionality of a region is enhanced or hidden, major changes can occur. Such processes therefore function, not only at the level of the cell, but in the organism as a whole. While some epigenetic changes are programmed (as in creating liver cells), others draw on an individual’s lived experience. Thus, in identical twins, the pattern of DNA expression is similar early in development. However, across the lifetime, a cascading set of epigenetic effects will draw on processes such as DNA methylation.

In other cases, genetic material remembers its maternal or paternal origin. This leads to manifestations in the overall likeness of an individual and is especially well known in so-called genomic imprinting. In mammals, all females are genetic “chimeras” because, in their cells, only one (of two) X chromosomes functions. In a given cell lineage whether this is maternal or paternal is determined at random. If the active chromosome bears a debilitating mutation, the effect cannot be mended in spite of the second (but inactivated) X chromosome has the right gene. Serious mental diseases may develop when maternal/paternal imprinting gets erased or impaired (e.g., Prader-Willi or Angelmann syndromes). In some groups (plants, and

---

[1] Such processes are especially important in the context of multicellular organisms and their ontogeny. It is important that some of them may outlive even to the next generation, thus transferring the experience of parents.

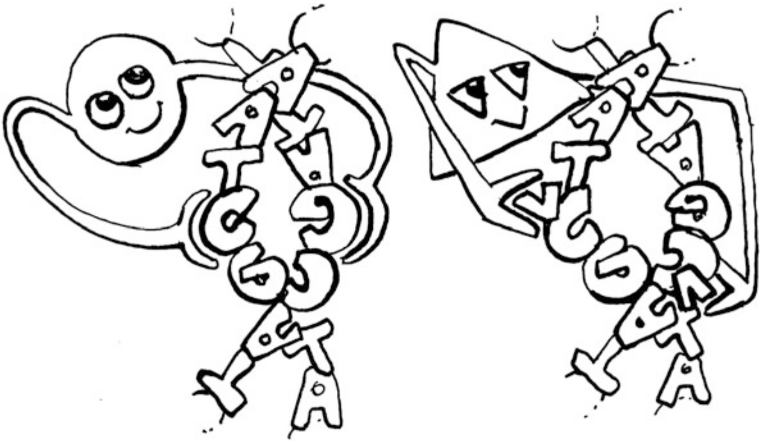


Fig. 10.5 Epigenetic marking: changing some characters affects the overall shape of a section on DNA. If the section AGCTAA represents a *ligand* for a specific regulatory protein (**a**), a modification (to AGČTAA) turns it into *another ligand*; it becomes the target of a protein (**b**). The complex DNA-protein participates in the cell's protein network by influencing its ability to read other parts of the DNA script: the "reading machinery" behaves differently in cases **a** and **b**

perhaps also animals), imprinting enables parents to transmit information to their offspring about the environment they are likely to encounter (e.g., Gilbert and Epel, 2009; Allis et al. 2009).

Nucleosomes

DNA strings (billions of "letters"—in mammals) are, in eukaryotic—animal, plant, fungal—cells organized into structures of higher order called *chromatin*: its lowest level of structuration is a "rosary" of *nucleosomes* containing about 147 DNA "characters" wrapped around 8 proteins (doublets of 4 different *histones*, see Fig. 10.6). While stabilizing the strand of DNA, these also enable or deny proteins access to particular sections of genetic material. This depends on functions that are independent of central control. Rather the actions of specific proteins (e.g., methylation, phosphorylation, acetylation), give rise to modifications (and erasures) of histone proteins whose end tails stick out from the nucleosome (e.g., Allis et al. 2009). The modified surface of the nucleosome can thus serve as binding site for proteins that constitute a chromatin ecosystem. Furthermore, such a modification affects all other proteins. It results in a network of interactions that maintains cell differentiation (e.g., as liver cells or neurons) while favouring quick and reversible response to external or internal cues. For example, some genetic material becomes walled up in a given cell lineage or during a developmental stage. By modifying both the DNA and histones, that part of the chromatin acts as an attractor that

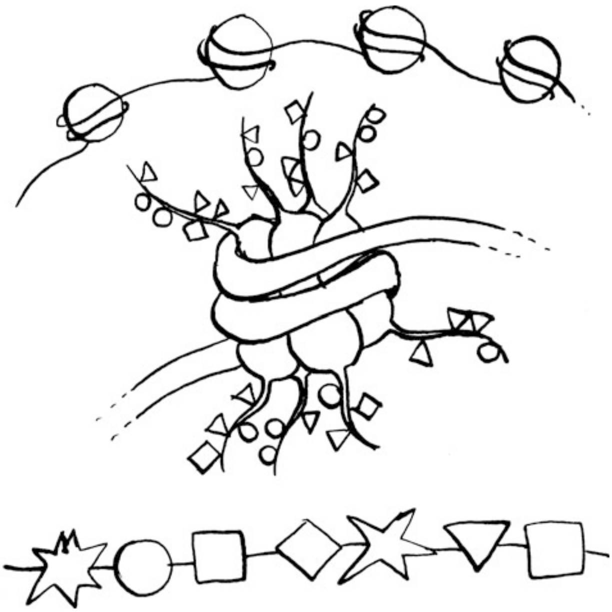


Fig. 10.6 The nucleosome. **a.** DNA is wrapped around 4 kinds of histone proteins. **b.** Histones are prone to binding by regulatory proteins; epigenetic marking (symbols on protruding "tails") can change the set of proteins that bind to a particular part of a histone. Such a change may switch the whole protein network into a different setting. **c.** Each nucleosome (plus proteins attached to it) thus represents a unique, fine-tuned complex that decides how and when the genetic script at that position is to be read. (After Allis et al. 2009)

silences part of the DNA string—possibly thousands of nucleosomes in a row. In other cases, protein assemblies organize regions to produce a given cell lineage. In most cases, even long-lasting modification may (or should) be reversible in circumstances such as regeneration or, gametogenesis. This view of the cellular ecosystem as akin to reading is shown in the nucleosome pictured in Fig. 10.6.

Elsewhere Markoš and Švorcová (2009) draw an analogy to a natural language that emerges in a natural community of living protein players ("speakers"). This, we argue, cannot be reduced to a fixed code that depends on a program being executed. The parallel is striking: while a histone code can be described in terms of (grammatical) rules, it draws on a dynamical, experience-dependent ecosystem or, simply, the total protein milieu. It is argued that any formal language defined as a set of character strings and determinate operations (Searls 2002) is merely derivative of natural language, that is, it was created by individuals (proteins, cells or humans) who live in the natural world. Developing a consensus on how to read these codes is historical and based on the experience of a community of natural speakers: as Love (2004) suggests, it consists in second-order constructs. Although rules can be described by formal languages, these do not *constitute* natural languages. Just as there are no transcendental laws or rules of human language, biological codes are unlikely to depend on a deeper formal language. Rather, just as in human languaging, biological meaning is extracted by natural 'speakers' who dwell in a historical world of bodily experience.

If the correlation between the DNA script and the shape of the protein is contextual, and experience dependent, then emancipation from the genetic script is likely to go further at "higher", supramolecular levels. Accordingly, we now trace parallels between the interactions of biological systems and the metabolic and symbolic aspects of language and, beyond that, what are usually regarded as different language-systems.

Symbiotic Interactions

In biology, there is often intimate coexistence between two or more lineages of organisms (Sapp 1994, 2003). This *symbiosis* includes endosymbionts that have been long established within the cells (e.g., the mitochondria or chloroplasts that are viewed as integral to eukaryotic cells), ones living inside other bodies (e.g., bacterial communities in bowels) and the more floating interactions that constitute ecosystems. Symbioses are ubiquitous: they serve the biosphere in that, for example, symbiotic bacteria perform activities that their hosts require. They manage photosynthesis, sulphur metabolism, nitrogen fixation, cellulose digestion, and the production of nutrients (e.g., Hoffmeister and Martin 2003). Symbiosis is thus mainly understood as persistent mutualism or, as "associations between different species from which all participating organisms benefit." Symbiotic interactions are not marginal, academic topic but, rather, resemble the distributed cognitive systems that allow humans to use artifacts and institutions to extend their cognitive powers. In the terms proposed by Douglas (2010): "Plants and animals live in a microbial world. *Their surfaces are colonized by microorganisms (bacteria and protists) from which they generally derive no substantial harm. Some plants and animals, however, live in specific and coevolved relationships with particular microorganisms, and* these associations have profound impacts on the ecology and evolution *of the taxa involved and, in some instances, also on entire ecosystems. In particular, animal or plants* symbioses *with microorganisms* dominate most terrestrial landscapes, certain coastal environments and the immediate environs of deep-sea hydrothermal vents. […] *Symbioses are important not just because they are widespread and abundant, but also because the acquisition of* symbiosis can dramatically alter the evolutionary history *of some lineages and change the structure of ecological communities.*" (Douglas 2010, pp. 19–23, emphasis added).

Although symbiosis can be compared with many aspects of human cognition, we focus on its ecological and evolutionary consequences. As an ecological force, symbiosis ensures that species are bound to cohabit. For example, terrestrial plants typically have an intimate symbiotic connection between their roots and fungi. The most ancient and widespread partnership is *arbuscular mycorrhiza* that dates back

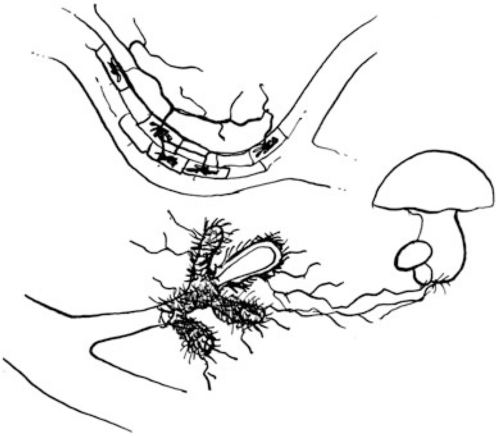


Fig. 10.7 Mycorrhizal symbiosis—tight cohabitation of fungal mycelium with roots of most plants. Two of many possible configurations are shown: **a.** Endomycorrhiza—fine mycelial protuberances invade the plant-cell cytoplasm and create an elaborated network. **b.** Ectomycorrhiza—while also very intimate, hyphae do not invade the interior of cells. The fungus interconnects trees within its reach, i.e. the whole forest may be networked in this way, the network involving many species of plants, fungi, and other organisms like bacteria

circa 460 million years and applies to 250 000 living plant species (Redecker et al. 2000; see Fig. 10.7). Fungi benefit plants by mobilizing nutrients from organic substrates while also delivering water. This is because fungal hyphae are thinner and thus permeate soil better than root hairs. In return, plants subsidize fungi by organic matter.

Symbiosis influences biological evolution profoundly. For example, new lineages of organisms can be engendered by the fusion of previously symbiotically living systems. Symbiogenesis is thought to have given rise to eukaryotic cells that draw on a conglomerate of different bacterial partners (see theory of serial endosymbiosis by Margulis 1993; Margulis and Sagan 2002). Indeed, even those who posit that nature is controlled by something like fixed codes admit that (at least) two kinds of cell organelles—mitochondria and plastids—originated from free-living microbial ancestors. (Douglas 2010; Margulis and Fester 1991; Overmann 2006; Paracer and Ahmadjian 2000; Sapp 1994, 2003). What is remarkable on symbioses is not the fact that different beings, like Russian dolls, share a composite body. Rather, what matters is that, unlike Russian dolls who are indifferent to each other, symbiosis involves mutual understanding between partners who spent even billions of years as separate lineages.

The moral of the story is becoming clear. In biology, wherever we look, we find interactive communities that "somehow" modify what first seems simple. Once we look "below the skin" of a cell, we find an ecosystem of cellular proteins that bend, prune, decorate and tattoo (but also clear away) other proteins: their existence is dependent on a genetic script but their fate depends on the field beyond. The same pattern appears at other levels: although all genes are present in every cell, their expression is distributed through the workings of structures and processes that will put down epigenetic markings. The unpredictability of the outcomes, that is the history of evolution comes to the fore when unrelated lineages enter intimate cohabitations. The same picture applies to ontogenies (i.e., patterning multicellular bodies). Development of a multicellular individual is a fascinating process especially when we trace its historical dimension across lineages and begin to consider what the biosphere has to say about such essentially intimate process.

Ontogeny

Many who discuss evolution echo the central dogma in claiming that *the potential of a species to evolve new traits is constrained by its genome or the set of genes it has available*. For example, Poe writes: "It might be evolutionary advantageous for your progeny to have wings, but it's simply not possible given the genes H. sapiens has to work with" (Poe 2011, p. 8). Whatever the truth of the claim its evolutionary basis cannot be what lies in the genome. Indeed, such a view is the biological counterpart of "written language bias" in linguistics (Linell 2005). Just as written letter strings are sometimes seen as basic, even primary, forms of language, DNA strings can be viewed that way. Function is ascribed to static, reproducible, and rational entities that can be seen and known in totality. Written language bias influenced molecular biology in the 1950s and 1960s (see Markoš and Faltýnek 2011) and, even today, some regard "linear biology" as biological common sense. Just as texts can be reduced to sequences (successions) of letters, DNA conforms to sequences of bases in nucleic acids and proteins. On this view, formal syntax lies 'behind' living phenomena—both language and the likenesses of living bodies. Indeed, the "central dogma" takes the extreme view that information is never ambiguous and flows from a script to the body.

The evidence presented above shows why we reject linear models in biology. First, simple proteins do not derive unequivocal shapes from nucleotides sequences. Second, distributed knowledge contextualizes script by assembling cells whose histories contribute to different lineages and organs. Third, members of different lineages use context to construct a world where cohabitation is widespread. Perhaps, then, we should return to our claim that Ezequiel's meaning cannot be extracted *solely* from a sequence of letters. In denying peace to the wicked (if that was his aim), the likeness (of a message, or of a body) is not a function of a sequence, program, or algorithm. Rather, it draws on a context that belongs to a given lineage, group, organism, and often does so creatively. Members of different species (≡ cultures) treat identical (or very similar) scripts ways that are quite specific: understanding a text is not a passive crystallization or decoding.

Vertebrates, arthropods, earthworms and even echinoderms have lost the two-sided symmetry of their ancestors. In the evolution of these *Bilateria*, all species have the same basic body plan (antero-posterior and dorso-ventral axes, left-right symmetry; see Švorcová 2012): differences arise from localized expression of ancient, conservative genes. The body plan is set by embryogenesis long before the appearance of body parts. Since bilaterian phyla have evolved independently for more than 500 million years, it is striking that the basic script remains unchanged. While the genes in each lineage underwent changes in "spelling" as some were duplicated, others deleted or otherwise modified, even unrelated lineages have much in common. For example, deletion of a single gene in the genome of fruit fly can be deleterious or lethal; however, the consequence can be experimentally reversed—by transferring a gene from the genome of a mouse (Gehring 1999). Although proteins coded by such homologous genes differ in many parameters, the message is 'understood': the fly embryo steers the homologue towards a normal developmental pathway. And, of course, "normal" is interpreted as flies (not mice). Thus, if one deletes the gene that initializes eye development in the fly embryo, blind flies will be born. However, a mouse gene restores the development of eyes-those of an insect not a mammal. Thus, a particular protein serves as a tool for establishing a developmental pathway: it does not determine the end product (the eye). Plainly the digital representation of genes (an inscriptional form that may be shared by fruit flies, mice and humans) does not determine how genes work. Rather, this is understood in the "cultural" context of the lineage (species, culture: at the lowest level of description, it depends on an embryo that grows in an ecosystem of interacting proteins—cells and tissues). This complexity allows the same genes to be used in many ways while nonetheless preserving (and transferring) the essentials of the proteins involved. The resulting patterns, ontogenetic outcomes, depend on bodily or lineage memory (see below), not on a linear string that enshrines a memory in a store or depot.

Just as in the Biblical story, the genes are written in an ancient script that is open to non- arbitrary interpretations. Understanding depends on both the individual *and* how the outcome is settled in a given population. The results depend on both situated and non-local factors. To illustrate this matter, one might consider the notorious comparisons between chimp and human genes. While now widely known that their genomes are 98% "similar", there is debate what such a number means (see, e.g., Marks 2002). Our comparison with reading mechanisms of book of life can be further elucidated by examples of inscriptions: thus an ancient philosopher's name is rendered as "Aristotelés" in Czech and "Aristotle" in English.[2] Is his message different for both communities? If it is not, as some will argue, this depends on the history of individuals and populations—not spelling.

Examples such as these may appear trivial. However, we should not assume that, in both life and culture, small changes can have large effects. Changing even a fraction of a percent of genetic material can make a difference—especially if the mutation affects a genomic control centre (Davidson 2006; Carroll 2005). In

---

[2] Versions of written US and UK written English may differ in the spelling in 2% of strings. Does this explain the differences between two nations? Remarkably, this line of thinking is pursued by those who seek a genetic Word (in DNA) that is "responsible" for differences in the appearance of the living being (phenotype).

presenting our case, we show only that it is naïve to posit the existence of virtual body plans that are attained (and perhaps even foreseen) by a single "keystroke", or a mutation that creates a "hopeful monster".[3]

The High Level of Meaning

We approach the most speculative part of our paper. Rappaport (2010, p. 10) argues: "The survival of any population, animal or human, depends upon social interactions characterized by some minimum degree of orderliness, but orderliness in social systems depends, in turn, upon communication which must meet some minimum standard of reliability if the recipients of the message are to be willing to accept the information they receive as sufficiently reliable to depend upon." We came to similar conclusions earlier when we compared the coherence of members of a biological species to a culture (Markoš 2002). Yet, Rappaport goes even further: his "standard of reliability" lies in rituals shared (albeit not always necessarily respected) by all living members of the community: it is the tie that defines it. Ritual, for him, is "The performance of more or less invariant sequences of formal acts and utterances not entirely encoded by the performers" (p. 24). In other words, it sculpts the "fashion" according to which "we" behave, even if there is no logical necessity to perform exactly in such a way, but it "constructs" the present, as well as the eternity of a given community. "Societies must establish at least some conventions in a manner which protects them from the erosion with which ordinary usage—daily practice—continuously threatens them." (Rappaport 2010, p. 323) "Universal sacred postulates" in rituals serve as such eternal constant that are not to be questioned, not even interpreted in various ways. Yet, they have their evolution across generations. May it be that biological species also constitute such a community kept together by the ritual inherited from the predecessors? Even if rituals seem eternal, they change in subsequent generations as the umwelt or "worldview" of a given lineage shifts in this or that direction. With a very similar "sacred texts", that is, the genome, we have—after 8 millions years of separation—two cultures of humans and chimps. If the parallel between languaging and life should be fruitful, we should be prepared to think in similar lines. How we look like today is the matter of our genes *and* of the ways how we make use of them in the ever-changing world.

---

[3] In European history, a single "mutation"—insertion of word *filioque* into the Christian creed (and Son) in the 6th century—is often seen as the main "cause" of schism between Orthodox and Western Christianity.

Conclusion

Life cannot be subsumed under physico-chemical principles (even expressible through mathematical notation) because, as Simon (1971) argues, biology and physical science have different objects. Simply said, physical systems lack *meaning*. The fact was first recognized in systems theory and cybernetics (e.g., Bertalanffy 1968); however, no scientific concept of meaning has been developed or needed by the exact and empirical sciences. It is possible, of course, that this is logically impossible or that it just cannot be achieved in quantifiable ways. However, organisms are both ontological and historical: they are products of phylogenetic and evolutionary history. Not only is their multi-scalar nature likely to contribute to the complexity of meaning but this is likely to depend on how relationships use hereditary material to develop over time. As we have seen, this depends on the spatial conformations of DNA molecules and interrelations between them (e.g., DNA–RNA, RNA-protein, protein–protein) that gives the living world has a character of a network or a web of interactions. To grasp the 'core' properties of biological entities, we always need to know about their exact setting. Conversely, it is *far from enough to rely on knowledge of the structure* of their elements. In developing the language metaphor of life (Markoš and Faltýnek 2011; Markoš et al. 2009; Kleisner and Markoš 2005, 2009), we challenge the view that only the lowest level of meaning is accessible to science. Rather, we examine higher levels, where "meaning" gradually becomes applicable to the realm of living. It is our view that this is the most appropriate basis for explaining life and placing it in a coherent system of knowledge that also gives weight to the complexity of human worlds that unfold within a cultural meshwork.

Acknowledgements Supported by the Charles University Project UNCE 204004 (AM), and by the Charles university grant 107/43-253166 (JL). We thank Stephen Cowley for innumerable comments, and for putting the text into tolerable English. Fatima Cvrčková is the cartoons' creator.

References

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2007). *Molecular biology of the cell*. New York: Garland Science.

Allis, C. D., Jenuwein, T., & Reinberg, D. (2009). *Epigenetics*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. The Nobel lecture. *Science, 181,* 223–230.

Barbieri, M. (2008a). Life is semiosis: The biosemiotic view of nature. *Cosmos and History: The Journal of Natural and Social Philosophy, 4,* 29–51.

Barbieri, M. (2008b). The code model of semiosis: The first steps toward a scientific biosemiotics. *The American Journal of Semiotics, 24,* 23–37.

Bergman, A., & Siegal, M. L. (2003). Evolutionary capacitance as a general feature of complex gene networks. *Nature, 424,* 549–552.

Bertalanffy, L. (1968). *General system theory: Foundations, development, applications*. New York: George Braziller.

Carroll, S. B. (2005). *Endless forms most improbable. The new science of evo devo and the making of the animal kingdom*. New York: Norton.

Cowley, S. J., & Vallée-Tourangeau, F. (2010). Thinking in action. *AI & SOCIETY, 25,* 469–475.

Davidson, E. H. (2006). *Genomic regulatory systems. Development and evolution*. San Diego CA: Academic Press.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dawkins, R. (1982). *The extended phenotype*. London: Freeman.

Douglas, A. E. (2010). *The symbiotic habit*. Princeton NJ: Princeton University Press.

Eco, U. (1994). *The limits of interpretation*. Bloomington, IN: Indiana University Press.

Gehring, W. (1999). *Master control genes in development and evolution: The homeobox story*. New Haven, CT: Yale University Press.

Gilbert, S. F., & Epel, D. (2009). *Ecological developmental biology: Integrating epigenetics, medicine, and evolution*. Sunderland, MA: Sinauer Associates.

Good, M. C., Zalatan, J. G., & Lim, W. A. (2011). Scaffold proteins: Hubs for controlling the flow of cellular information. *Science, 332,* 680–686.

Hall, B. K., Pearson, R. D., & Müller, G. (Eds.). (2004). *Environment, development and evolution. Toward a synthesis*. Cambridge, MA: MIT Press.

Heidegger, M. (1959/1982). The nature of language. In *On the way to language* (pp. 57–108). San Francisco CA: Harper.

Heidegger, M. (1995). *The fundamental concepts of metaphysics. World, finitude, solitude.* Bloomington, IN: Indiana University Press.

Ho, M.-W. (1993). *Rainbow and the worm*. Singapore: World Scientific.

Hoffmeister, M., & Martin, W. (2003). Interspecific evolution: Microbial symbiosis, endosymbiosis and gene transfer. *Environmental Microbiology, 5,* 641–649.

Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An eternal golden braid*. New York: Basic Books.

Kauffman, S. A. (2000). *Investigations*. Oxford: Oxford University Press.

Kirsh, D. (2010). Thinking with external representations. *AI & SOCIETY, 25,* 441–454.

Kleisner, K., & Markoš, A. (2005). Semetic rings: Towards the new concept of mimetic resemblances. *Theory in Biosciences, 123,* 209–222.

Kleisner, K., & Markoš, A. (2009). Mutual understanding and misunderstanding in biological organisms mediated by self-representational meaning of organisms. *Sign Systems Studies, 31,* 299–309.

Linell, P. (2005). The written language bias in linguistics. London and New York: Routledge.

Love, N. (2004). Cognition and the language myth. *Language Sciences, 26,* 525–544.

Margulis, L., & Sagan, D. (2002). *Acquiring genomes: A theory of the origins of species*. New York: Basic books.

Margulis, L., & Fester, R. (Eds.). (1991). *Symbiosis as a source of evolutionary innovation. speciation and morphogenesis*. Cambridge MA: MIT Press.

Margulis, L. (1993). *Symbiosis in cell evolution*. New York: W. H. Freeman and Company.

Markoš, A., & Faltýnek, D. (2011). Language metaphors of life. *Biosemiotics, 4,* 171–200.

Markoš, A., & Švorcová, J. (2009). Recorded versus organic memory. *Biosemiotics, 2,* 131–149.

Markoš, A. (2002). *Readers of the book of life*. New York: Oxford University Press.

Markoš, A., Grygar, F., Hajnal, L., Kleisner, K., Kratochvíl, Z., & Neubauer, Z. (2009). *Life as its own designer: Darwin's origin and western thought*. Dodrecht: Springer.

Marks, J. (2002). *What it means to be 98% chimpanzee: Apes, people, and their genes*. Berkeley, CA: University of California Press.

Matsuda, R. (1987). *Animal evolution in changing environment with special reference to abnormal metamorphosis*. New York: Wiley & Sons.

Monod, J. (1972). *Chance and necessity*. Glasgow: W. Collins Sons.

Overmann, J. (Ed.). (2006). *Molecular basis of symbiosis*. Berlin: Springer.

Paracer, S., & Ahmadjian, V. (2000). *Symbiosis. An introduction to biological associations*. Oxford: Oxford University Press.

Poe, M. T. (2011). *A history of communications: Media and society from the evolution of speech to the internet*. Cambridge: Cambridge University Press.

Rappaport, R. A. (2010). *Ritual and religion in the making of humanity*. Cambridge: Cambridge University Press.

Redecker, D., Kodner, R., & Graham, L. E. (2000). Glomalean fungi from the Ordovician. *Science, 289,* 1920–1921.

Rutherford, S. L., & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature, 3*(96), 336–342.

Sapp, J. (1994). *Evolution by association: A history of symbiosis*. Oxford: Oxford University Press.

Sapp, J. (2003). *Genesis: The evolution of biology*. New York: Oxford University Press.

Searls, D. B. (2002). The language of genes. *Nature, 420,* 211–217.

Simon, M. A. (1971). *The matter of life: Philosophical problems of biology*. New Haven, CT: Yale University Press.

Švorcová, J. (2012). The phylotypic stage as a boundary of modular memory: Non mechanistic perspective. *Theory in Biosciences, 131,* 31–42.

Taipale, M., Jarosz, D. F., & Lindquist, S. (2010). HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nature Reviews Molecular Cell Biology, 11,* 515–528.

Thibault, P. J. (2011). First-order languaging dynamics and second-order language: The distributed language view. *Ecological Psychology, 23,* 210–245.

Waddington, C. H. (1975). *The evolution of an evolutionist*. Edinburgh: Edinburgh University Press.

Chapter 11
Cognition and the City: Cognitive Ecology and the Paris Commune of 1871

Alexander Aston

Abstract By investigating flows of energy, matterand information during the Siege and Commune of Paris from 1870 to 1871 this analysis attempts to show how human cognition intersects with its environment to form self-organizing, complex adaptive systems. While traditional explanations of the Commune have generally revolved around Marxist analysis, it is possible to analyse the events of 1871 as transformations in the urban and cognitive ecology of the city. Specifically, utilizing the perspectives of Material Engagement Theory and distributed cognition, this paper explores the ways in which cultural materials feedback into cognitive processes to shape social activity. Changes in Paris' urban ecology produced selection mechanisms which facilitated Parisians organising around different institutional settings. Radical clubs, vigilance committees and worker's cooperatives were able to provide stability in a rapidly degrading urban environment, and thus a point of departure for new forms of social organisation to emerge. To facilitate this process, Parisians modified their environments, both within and outside of these institutional settings, in ways that altered the flow of information through the city and provided new ways of engaging in a revolutionary context. Parisians utilised material artefacts such as rifles, flags and bodily decorations in order to distribute cognition, enabling collective revolutionary action. This paper shows that the most important feature of urban environments is the ability to facilitate individual and collective adaptation to ecologies dominated by the physical and cognitive presence of their own species. In this view, cities are understood as selection driven adaptive landscapes, co-evolutionary structures that emerge to facilitate and sustain dense human habitation through the material organization of cognition. The transformations in Paris' urban ecology led to Parisians reconfiguring their cognitive environments, precipitating the development of radical social institutions. Artefacts, circulating in human ecologies, function to entangle human cognition and behaviour into coherent environmental relationships. Thus, human societies do not fundamentally break from the natural world but are part of a developmental continuum with evolutionary and ecological dynamics.

A. Aston (✉)
School of Archeology, Oxford University, Oxford, UK
e-mail: alexander.aston@keble.ox.ac.uk

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_11

Introduction

On August 31 1870, the French military began strategically demolishing houses in the suburbs of Paris. The refugees flooding into the city appeared to Edmund de Goncourt as “the migration of an ancient people” (de Goncourt et al. 1962, p. 168). Behind this inundation of bodies, more than a million Prussian soldiers poured across the borders of France. In little over a month, the Germans had employed rail, telegraph and artillery with a devastating synergy to overwhelm the armies of France’s Second Empire. With increasing waves of shock, it became apparent that Paris needed to prepare for a siege. Accustomed to dominance, the high command’s response proved haphazard and confused. Citizens wandered in a daze as ragged and weary conscripts set up camp on the Champs de Mars; soldiers erected barricades around the Arc de Triumph. Strategic locations such as Châtillon were abandoned and civilians that should have been evacuated were brought into the city walls. Military officials hastily cleared areas to store livestock and produce, but they devised no system of rationing. Whatever could not be brought into the city, officials burned in the fields, a scorched earth policy to deny the Prussians access to provisions. Paris was in the early stages of a profound social transformation. For American ambassador Elihu Washburne, the city was “one big camp… Streets and avenues are filled with tents and baggage wagons, horses, forage & c. The garden of the Tuileries is filled with artillery.” (Washburne and Michael 2012, p. 50). Over the next four months, with over a million and a half people confined and isolated within its walls, the “City of Light” went dark, the grand avenues denuded of trees and sewers scoured of rats. The Commune emerged in the wake of this devastation.

The Prussian encirclement produced a considerable bottleneck in the flow of matter, energy and information through the city, transforming its urban ecology. The relationships that sustained Parisians altered in terms of both food, fuel and the positions of different communities in relation to those sources. During this process the lower and underprivileged classes were simply priced out of the food supply. The only major source of income during the siege, the National Guard, paid one and a half Francs a day. By the middle of October, a single egg cost one franc (de Goncourt et al. 1962). Towards the end of the siege the death rate more than tripled to 4,670 a week, mostly the children and elderly among the poor (de Boismont 1871). Under these pressures Paris began to undergo something akin to social “speciation” between the classes. The German siege forced working class members of Parisian society into marginal social niches while the more elite members were able to tap exotic and exorbitant resources such as the zoological gardens. As the Urban metabolism of the city began to shut down individuals and communities began to radically alter their environments, scouring Paris for food and fuel. By January, the city was nearly denuded of trees; the streets and buildings were unlit for want of gas, no horses or carriages moved about the avenues. Few humans braved the streets other than hordes of the destitute scouring the city for wood to burn. The effects of the drastic reconfiguration of energy-matter produced an institutional selection mechanism. For example, the lack of fuel helped bolster political club attendance. In one of the coldest winters ever recorded radical clubs became attractive environments simply for warmth. Critically, the lack of fuel also shut down less radical, more energy intensive alternatives to the clubs. The scarcity of fuel made the staging of theatre productions increasingly difficult, resulting in their total closure by January. With the closure of the theatres attendance at the clubs rapidly increased, especially among poorer communities. The contemporary social theorist Gustave Molinari lamented, "la fermeture de theatres, accordait une veritable prime d'encouragement a la formation des clubs" (de Molinari 1871, p. 6). The effects of the siege helped to push working class communities into the radicalized environments that provided literal insulation from the cold. As the adaptive landscape of the city began to shift, Parisians responded to these changes by transforming their environments in ways that afforded new possibilities of thought and action. Individuals and communities utilised a scaffolding of radical culture such as Jacobin clubs and socialist cooperatives as well as employment in the national guard to provide stability. This emphasized new networks of cooperation which emerged as the former organisation and connectivity of the city began to break down.

Following a brief armistice, in March of 1871, with a monarchist coup d'état appearing imminent, Parisian communities rose up in the largest urban insurrection of nineteenth-century Europe. After nearly five months of freezing and starving through siege, a multitude of local associations and groups operated Paris from the ground up for seventy-two days. Parisians formed workers' cooperatives, neighbourhood associations and women's unions, creating their own municipal infrastructure to meet the needs of their communities. Highly democratic elections brought a working class population into the political process. The Commune produced and distributed goods, minted coins, established schools, medical facilities and newspapers, passed social measures, collectivized factories, and generally maintained the city. Paris quickly developed a relatively self-sustaining and radically new pattern of social organisation. Five days after officials declared the start of the Commune, fighting broke out with the National Assembly at Versailles. For over two months, the organizations that emerged out of the rebellion supported Paris until Versaillais troops gained access to the city's streets. For the seven days of la Semaine Sanglante (the Bloody Week) Parisians and soldiers spilled blood across barricades and boulevards as the city burned. In the aftermath, soldiers executed as many as 30,000 in the streets during one of nineteenth-century Europe's largest massacres. Later, police condemned to death, imprisonment or deportation nearly 13,450 participants of the Commune (Shafer 2005). In doing so, officials crushed one of the most extensive and radical social experiments in history.

The Bee and the Architect

The field of history has long demarcated its academic territory with the development and spread of the written word (Smail 2008). The empirical emphasis on recorded language is where historians have drawn the disciplinary boundaries around forces considered to be internal and external to the historical process. It is argued that language reveals the intentionality of the human as a historical agent in juxtaposition to the blind causality of a mechanistic non-human world (Collingwood 1951). Historical causation is seen to arise from human agency imposing a symbolic reality upon an external and relatively passive material substrate. Thus, things such as the non-human environment, material culture, and the body have a tendency to be analysed through linguistic models and metaphors such as the post-structural emphasis on representation and discourse (Foucault 2005). These theoretical assumptions issue from the conceptual dichotomies of mind and matter, humanity and nature that are endemic to western epistemology. There is, of course, a logic to these dualisms, the properties of human cognition are not easily accommodated by the mechanisms described in the physical and biological sciences. Yet, these very sciences increasingly tell us that these processes are interdependent, that human beings are developmentally embedded in a world of non-linear and emergent dynamics in which causation is not easily disentangled. Increasingly, historians are realising that their analyses must accommodate causal forces far beyond limited notions of human agency and intentionality as the engines of historical change (Chakrabarty 2009). The evolution of human beings has introduced a novel ecological force to the planet and historical models need to account for these dynamics if the idea of human agency is to have any substance. The unique features of human cognition and behaviour are able to interact at a vast number of scales which enable and constrain historical possibility. Indeed, we are learning that human activity is capable of altering flows of energy and matter on planetary and geological scales. This is a kind of agency not easily accommodated by linguistic models of historical causality. Rather, historical analysis should seek to approach cognition and the material environment not as distinct realms of causation but as integrated developmental process through which the contingencies of history emerge.

Marx claimed the Commune was a harbinger of the Proletarian revolution and his theories remain the most common analytical framework for understanding the uprising. However, the Paris Commune has never been explored as an inherently ecological phenomenon, an organic transformation of a society emerging from the reorganization of Paris' material, cognitive environments. Marx wrote in Capital that, "A spider conducts operations which resemble those of the weaver, and a bee would put many a human architect to shame by the construction of its honeycomb cells. But what distinguishes the worst architect from the best bees is that the architect builds the cell in his imagination before he constructs it in wax" (Marx and Ben 1990, p. 284). Yet, akin to the beehive, the city can be understood as a powerful example of ecosystem engineering and niche construction. Urban environments are entangled with the evolution of the human body, for the structure of a city to survive it must be able to sustain physiologies with a degree of adequacy. What makes this unique in relation to other niche constructing organisms is that human beings utilise material culture to structure and distribute cognitive abilities and tasks in ways that facilitate cooperative behaviour and self-organising dynamics at multiple scales. Rather than structure being pre-imagined in the mind and imposed upon the world, what becomes apparent is an active process through which structure emerges. In these regards new approaches to the relationship between cognition and the environment potentially have much to offer historiography. Over the past three decades the development of *Enactivist*, *Embodied, Distributed* and *Extended* models have sought to analyse the environmentally situated feedback that shapes cognition (Varela et al. 1992; Hutchins 1995; Clark and Chalmers 1998). By investigating the mind as something that is developmentally embedded in the environment it is possible to explore how individual and collective actions are mediated through material things to produce historical phenomena. In these regards cities can be analysed as "cognitive ecologies" that form through processes of "cognitive developmental niche construction" (Hutchins 2010; Stotz 2010). This affords an opportunity to explore the entanglement of human cognition with its environment and assess historical change as an organic process while avoiding the pitfalls of reductionism and determinism.

If history is to be understood as an ecological process, then it is material culture that presents itself as the most obvious point of departure for such an analysis. Artefacts occupy something akin to the midway point between the human and the non-human. Simultaneously cognitive and environmental, they blur the boundaries of what is internal and external to the human form. Archaeologist Ian Hodder describes "things" as "nodes for the flows of inter-linked bundles of matter, energy and information" (Hodder 2012, p. 48). It is in these regards that the objects of human production are ecological phenomena, their very presence in the world altering environmental dynamics at a number of temporalities and scales, affording new interactive possibilities.[1] Recently, the development of Material Engagement Theory (MET) within the emerging field of cognitive archaeology, provides unique insights to the implications of material culture upon cognition (Renfrew 2001; Malafouris 2013). MET explores 'the synergistic process by which, out of brains, bodies, and things, mind emerges' (Malafouris 2013, p. 17). Specifically, material engagement emphasises the "metaplastic" feedback created by neural, behavioural and material plasticity (Malafouris 2013). In this view, material culture provides the medium through which human individuals and collectivities are able to reconfigure their environments in ways that afford new cognitive and behavioural opportunities. The flows of energy, matter and information in which human beings are immersed are partially structured and made available for manipulation through material culture. This enables individuals and societies to shape their developmental dynamics through the spatial and temporal organisation of material relationships. Artefacts, acting as shared cognitive resources help to embed individual ontogeny within social, ecological and evolutionary processes. In this light, the cognitive ecology of the Commune was a "complex adaptive system" that emerged through sweeping changes in the flow of energy-matter through the city. This triggered selective

---

[1] The written word is a clear example of this. It transmutes the energetic patterns of sound into a new material and perceptual structure, losing a degree of informational complexity and dynamism in return for a slower rate of entropy in the environment and the ability for increased distribution.

pressures that helped restructure communities around new institutional frameworks and relationships. Parisians mediated these ecological dynamics by modifying their environments with things such as barricades and red flags, allowing them to conceptualize and collectively participate in a revolutionary context. It was from this drastic reconfiguration of Paris' ecology that a unique social form emerged for a brief moment in the spring of 1871.

Cognition in the City

Parisians proclaimed the Third Republic on September 4, 1870 after a series of defeats by the Prussians, Louis-Napoleon III's Second Empire collapsed. Edmund de Goncourt, witnessing the declaration of the republic at the Hotel de Ville, recalled "many were carrying branches in their hands and had green leaves fastened to their hats. There were a few soldiers with twigs tied to the barrels of their rifles." As the scene unfolded, "right at the top of the facade, a man tore the blue and white stripes from the tricolour, leaving only the red waving in the air" (de Goncourt et al. 1962). The red flag developed significance during the 1789 revolution; however, over the course of the nineteenth-century it had come to be identified with revolutionary socialism. After the suppression of labourers during the June uprising of 1848, the indefatigable revolutionary Louis Auguste Blanqui wrote, "The tricolour flag… has been twenty times bathed in the blood of the workers. The people raised the red colours on the barricades in '48, just as they raised them on those of June 1832, April 1834, and May 1839 … From this day on, these colours are theirs" (Blanqui 2011, p. 89). From strapping branches to rifles and ripping the flag, Parisians purposefully altered materials in the physical environment to facilitate new ways of conceptualizing the changing circumstances.

The significance of what the anonymous individual did atop the Hotel de Ville would not have been lost on anyone who saw it. The action of tearing the tricolour helps illuminate the plasticity of social space in shaping experience. The liberal leaders that situated themselves in control of the September 4 revolution wanted a moderate, political republic. However, the altering of the tricolour made clear that there were those present ready to demand the Social Republic. The red flag was an "enactive material sign" that facilitated Parisians thinking through their cognitive environments. The "material sign", argues cognitive archaeologist Lambros Malafouris, "does not primarily embody a communicative or representational logic but an enactive one. For material semiosis meaning is not a product of representation; it is a product of conceptual integration between material and conceptual domains" (Malafouris 2013, p. 18). The physical properties of flags, colourful and dynamic in the wind, attract the eye. The red flag was a way of projecting possibility as well as demarcating space, both in terms of territorial possession and contestation. Among the first things done when the Commune occupied a space was to raise the red flag. When the revolutionaries occupied the forts of the left bank on March 18, the citizens of Paris quickly saw the "red flag flying over the three fortresses" (Bellinger and Mitchell 2010, p. 134). The banner communicated the geography of revolution. Parisians were capable of reading the organizational state of their city, the territoriality of ideas, simply by scanning the skyline. "A common feature of mark making," Malafouris explains:

> Is that it is able to transform humans direct and immediate perceptual relation to the world. Such a shift in perception affords and stimulates new opportunities for enactive thinking. An interesting feature of this cognitive ecology of mark making is that it can leave persistent and visible material traces that alter the epistemic landscape of activity and thinking. (2013, p. 205)

The flags functioned as cognitive, epistemic artefacts, used "as both additional memory and as potent symbol-manipulating arenas" (Clark 2008, p. 12). The visual perception of the flags allowed for rapid information flows so that people could understand the overall condition of events such as military contestation.

Traditional accounts of the Commune tend to focus on elements such as the political tensions, economic forces, demoralization of soldiers and discourses that all played a role in triggering the insurrection of March 18. Commune historian Martin Philip Johnson has divides explanations of the Commune into one of two general "schools" composed of political and social interpretations (Johnson 1996). On the one hand, the political interpretation views the Commune as a response to the ravages of the Prussian siege and the insensitivities of the government, a position exemplified by Alistair Horne's (2007) *The Fall of Paris*. On the other hand, social interpretations have sought deeper patterns rooted in the Second Empire (1852–1870) and France's overall social and economic development. Structural Marxist accounts such as David Harvey's (2005) have reified the class-based narrative of the Commune's development from the flows and contradictions of capital during the Second Empire. Specifically, such analyses have focused on the implementation of the Haussmann Plan, which had reconstructed the core of the city. Roger V. Gould has rejected the class conflict account and argued that the Commune was essentially a conflict between the Parisian Community (and more specifically its neighborhoods) and the State on the issue of municipal liberties. For Gould the conflict emerged through organizational forms that favored specific "participation identities" or the ways in which individuals understand their roles and relationships in society through institutional settings (Gould 1995). While these are all important factors, they do not tell the whole story. The Commune was not a planned event, it was not executed by a revolutionary vanguard, rather it emerged through an ecology of individuals, communities, and institutions self-organising from the bottom up. In these regards, the events of 1871 bear more resemblance to Marx's bee than his architect. The Commune was imagined as it was made. Parisians responded to the ecological transformations of the siege by reconfiguring their environments, affording not only survival but new ways of thinking about social relationships. While political, economic and sociological accounts are of great importance, the Commune was shaped by deep patterns in the relationship between human cognitive evolution and the environment. If we wish to understand how coherent and sophisticated social structures form out of decentralised and distributed relationships, it is critical to explore the ecological properties of cognition and the ways in which material culture facilitates thought and action. Anyone who has ever tried to organize a protest of even a few dozen people will attest to the difficulty of coordinating elements such as marching routes, message and specific actions taken. The difficulty of organizing a revolution is another order of magnitude entirely. Yet, Parisians spontaneously overthrew their government without command from any centralized authority. In the case of the Commune, there were of course linguistic elements to the flow of information such as revolutionary posters, newspapers and speeches that helped organise communities. However, it is also critical to look at how cultural objects externalized "the processing of information" in non- representational ways. "The focus" argues Andy Clark, "shifts from accurately representing an environment to continuously engaging that environment with a body so as to stabilize appropriate co-ordinated patterns of behaviour." (Clark 2011, p. 16) It is through engagement and feedback with shared developmental affordances that humans are able to collectively self-organise.

During the attempted revolution of October 31, Edmund de Goncourt described the square at the Hotel de Ville as "a forest of rifle butts raised in the air" (de Goncourt et al. 1962, p. 177). Parisians were using the reversed rifles as cognitive "scaffolding" to communicate information in a distributed, non-representational manner. The habit of flipping rifle stocks was a common practice at the time for indicating peaceful intent as well as revolutionary behaviour. As material signs, the reversed rifles were "tagged," as Andy Clark explains, so that "complex properties and relations in the perceptual array are, in effect, artificially reconstituted as simple inspectable wholes. The effect is to reduce the descriptive complexity of the scene" (Clark 2011, p. 46). By reducing such complexity "cognitive artefacts" aid the rapid processing of large sets of information by individuals and groups enabling them to modify behaviour in appropriate and coordinated manners. The flipping of the rifle butts in the air allowed for quick visual communication of the overall state of the various armed groups, providing rapid informational feedback. From the perspective of the individual, seeing the reversed rifles would allow them to assess the degree of danger with a rough and ready calculation of the crowd's disposition. Clark describes the physiological process involved in such information gathering, "when our brains detect a sudden flash and our eyes automatically saccade in that direction, the motor routine embodies a kind of hard-wired implicit metacognitive commitment to the effect that we may gain useful, perhaps lifesaving, information by such a rapid saccade" (Clark 2011, p. 74). From this standpoint rifles do not represent violence, they embody it. In these regards, the rifles also functioned as externalised memory, their presence in social contexts was a form of enactive remembrance of past violence and future possibility. The binary manipulation of the rifles position added a dimension of communicative plasticity to the artefact. There is no clearer example of how this distributed cognition than the actual moment in which the insurrection first took hold during the 18th of March when the National Assembly sent soldiers to requisition the cannons created through public subscription and controlled by the Parisian National Guard.

Revolutionary Adaptations

It was dark when the military deployed to the Cannon Park atop the Butte of Montmartre. The National Assembly ordered the soldiers to take possession of the ordnance. As the sun rose over the city, people awoke to find that soldiers had overrun the National Guard checkpoints and controlled the artillery. Awareness of the situation grew with mounting intensity, bells rung and people streamed out of their homes; crowds of women, children, and the elderly, surrounding the soldiers, raised the alarm. Yelling at them, pleading, "the women of Paris covered the cannon with their bodies" (Michel and Nic 2004, p. 35). Soon National Guardsmen joined the tumult. Four times General Lecomte ordered the soldiers to fire on the crowd and four times the soldiers refused. Calling out to one another, National Guardsmen and soldiers began to reverse their rifles so that the stocks raised in the air. One anonymous eyewitness account from the events describes a clear positive feedback loop in the reversing of rifles among the soldiers:

> Rifles were placed on shoulders; muzzles of cannons were lowered. The crowd trembled but it did not budge. In a short but profound silence, the word resounded, "Fire!"
>
> The agony was piercing. The national guardsmen prepared to avenge the crowd if the troops fired.
>
> They refused to obey. One gun, then ten, then a hundred were turned up, and it seemed that the death that had hovered over this multitude took flight and spared them. (Anonymous quoted in Gulickson 1996, p. 42)

With stunning rapidity, the French regulars "surrendered their chassepots and even cried Vive la République! … the bulk of the soldiers went over to the people" (Lissagaray 2012, p. 66). The crowd surged forward and embraced the mutineers, disarming the officers and gendarmes before they could counter. It was the catalyst triggering an insurrection that cascaded through Paris. By the end of the day, the national Government ordered the immediate evacuation by all that obeyed and revolutionary Parisians controlled the city. It was a spontaneous and drastic reconfiguration of social power in Paris.

When the soldiers were attempting to seize the guns from the National Guard, the situation was intensely chaotic. A mass of angry, electrified people, some of them armed, shouting and yelling from all directions surrounded the armed men; bells, bugles and drums all sounded while the order to fire on the crowd rang out. Only a few months before, on January 22, soldiers had opened fire on a protest at the Hotel de Ville and killed some fifty civilians. However, the actions of soldiers and national guardsmen flipping rifle butts in the air helped reinforce the dynamics of rebellion. As cognitive artefacts, the rifles formed "material anchors" that grounded thinking processes in a tumultuous environment. Reversed rifles allowed citizens and soldiers to literally see each other's thoughts and feelings. Emboldened by the numbers that were refusing to fire the crowd began embracing and speaking with the soldiers. Bodies, sounds and artefacts created feedback that supported the conditions and possibilities of the situation. In the words of Malafouris, "the material physical qualities of artefacts do not depend on mental states but rather constitute those states." (Malafouris 2013, p. 164). In other words, by allowing Parisian to spatially process an insurrectionary moment the rifles formed part of the scaffolding on which the Commune was built. Shared developmental experiences of rifles ensured coherence in regards to their use in large groups. It is in this way that manipulation of these materials scaled up plasticity from individual behaviour into complex organisational dynamics. The rifles provided a visual binary that rendered the soldiers intentions as perceptual information, providing collective knowledge to the various individuals and groups within the mass. With the recognition that the soldiers did not intend to fire an extremely heterogeneous mix of peoples, defined primarily by opposition between civilians and the military, swiftly took on new behavioural dynamics. Acts of social solidarity, celebration and a shared antagonism towards military and government leadership spread from Montmartre through the neighbourhoods and districts of Paris. The various groups and institutions that had formed during the siege quickly aggregated around these dynamics, fuelling the insurrection. The rifles are a small, yet critical element of this process. They were a subtle catalyst in a phase transition that supplanted the National Government with the Commune. What this exposes are connections between material culture, the visual system, group behaviour and the transformation of social dynamics in an urban environment. A behavioural-cognitive shift in the conflict's social configuration was enacted through the rifles. This provided a core dynamic around which further revolutionary behaviour could aggregate and spread through the city enabling new social configurations to emerge.

In the tumult of insurrection, Parisian citizens of all perspectives were doing their best to interpret circumstances and adjust behaviour through what Clark describes as the "active self-structuring of the flow of information" (Clark 2011, p. 201). The presentation of bodies and materials became a critical component of communication and interpretation in the revolutionary environment. Early in April, Gustave Cluseret, head of the War Delegation, issued a communique: "I have observed with sorrow that, forgetting our modest origins, the ridiculous mania for braids, embroidery, and ribbons has begun to make its appearance amongst us." (Cluseret and Mitchell 2010, p. 22) Edmund de Goncourt wrote of "studying people's faces, which are a sort of barometer of events in revolutionary times" (de Goncourt et al. 1962, p. 185). Upon seeing a group of National Guardsmen, "wearing bunches of lilacs," he felt reassured that the Versaillais had defeated them because of the dejected looks on their faces. Later, Goncourt was discouraged to see Fort Issy "still flying the red flag" (ibid, 185). Facial features as well as uniforms and revolutionary accoutrement all became critical components for interpreting the condition of the city. This illuminates the synesthetic aspects of engaging with the environment. Ian Hodder defines synaesthesia "as a normal dimension of experience in the world in which sensory coherence is sought across domains" (Hodder 2012, p. 125). Sensory experience calibrated the individual to the condition of Paris, as is seen clearly in the case of sound. Elie Reculs wrote, "We hear the rolling of drums, and there they are coming from the depths of Saint Antoine quarter, coming down from Belleville and Montmartre, battalion after battalion. They've unfurled their red flags, they sing the 'Marseillaise,' they shout

"To Versailles, to Versailles!" (Reclus and Mitchell 2010, p. 98) Sonic information became a way of delineating spatial relationships. Drum rolls or cannon fire could identify specific areas and communities, giving an indication of their activity. Contextual knowledge, such as Bellville being a working class and insurrectionary community, could be used to interpret the meaning of the activity. In this sense, material engagement can be seen as a process through which collectives of humans regulate and self-organise their behaviour. Material culture provides an interactive medium through which individual and group behaviour can produce rapid, distributed feedback within larger collectivities to form coherent social phenomena with emergent properties.

Deeply embedded in the flows and fluxes of their environment, Communards began to modify the city in order to conceptualise and support their new social relationships.

Epistemic Engineering

Material elements such as guns and flags are only components of the cognitive scaffolding through which Parisians adapted to changing conditions. Paris itself provided the overall cognitive architecture. Steven Johnson describes cities as, "a kind of pattern amplifying machine: its neighbourhoods are a way of measuring and expressing the repeated behaviour of larger collectivities-capturing information about group behaviour, and sharing that information with the group. Because those patterns are fed back to the community, small shifts in behaviour can quickly escalate into larger movements" (Johnson 2001, p. 40). A city is, in a sense, the congealing of materials such as stone, timber and steel, into configurations capable of supporting human habitation. A key part of the way in which these materials support human habitation is how they support behaviour and cognitive processes. Malafouris argues that, "the distributed cognitive space is not simply the passive background against which the activity unfolds; it is something that can be used as a cognitive artefact" (Malafouris 2013, p. 67). Human beings organise their environments in order to support and streamline cognitive abilities in what David Kirsh describes as the "intelligent use of space." The physical manipulation and structuring of material relationships makes "it easier to stay in control of activity, we rely on techniques which reduce the memory load of tasks, the amount of internal computation necessary, or which simplify the visual search and categorization that is inevitably involved in performance." (Kirsh 1995, p. 65) In this sense, urban environments can be seen as nested relational structures in which individual engagement with artefacts link collective acts of engineering to shape the spatial and temporal organisation of the city. As Patterns of habitation emerged in Paris over the years, certain areas were dominated by craft workers or commercial interests, others areas came to be identified with particular cultures, behaviours, ideas and services such as the Bohemian left bank. Clark terms this "Cumulative downstream epistemic engineering" in which, "Groups of humans engineer their own habitats and these are transmitted to the next generation, who further modify the habitat. Importantly, some of these modifications are to the epistemic environment and affect the informational structures and opportunities presented to each subsequent generation." (Clark 2011, p. 66) Spaces within the city, modified by human habitation over time, were inscribed with certain qualities, and under specific condition, supported certain behaviours. There are few better examples in history of this phenomenon than the Hôtel de Ville.

The administrative seat of Paris was embedded into the organizational memory of the city as the nexus of revolution since 1789. Described as the "mecca of protest" in a brochure from the era, every successful revolution in France was proclaimed on the Hôtel de Ville's steps (Harrison 2000). Manuel de Landa argues that, "once the internal operations of an organization have become routinized, the routines themselves constitute a kind of 'organizational memory.' (de Landa 1997, p. 146) The Place de Grève, the square in front of the Hôtel where crowds gathered, had long been imprinted with particular social relationships such as public executions, migrant labour markets, and intense surveillance by authorities. Historian Casey Harrison explains the structural qualities that facilitated these relationships, "as the physical centre of the city and one of its largest open areas, the square was also a natural setting for the germination of crowds." (Harrison 2000, p. 412) The Hôtel de Ville, a location long imbued with the spectacle of political power and contested social space, provided vital cognitive scaffolding for engaging with revolutionary activity. When such conditions presented themselves large masses of peoples would gather in the square and demand a new government. In a sense, the physical space of the Hôtel de Ville functioned as a device for the storage of a collective revolutionary "memory" around which people could gather and "execute" a revolutionary "program." In the Hôtel de Ville, we can see what Clark describes as a "cognitive niche" where "physical structures combine with appropriate culturally transmitted practices to enhance problem solving and, in the most dramatic cases, to make possible whole new forms of thought and reason" (Clark 2011, p. 63). The material environments of Paris provided a critical means for shaping and supporting cognition. This can be seen clearly in the Parisian response to the brief Prussian occupation after the armistice. With incredible restraint, Parisians transformed their city to communicate their disdain for the German Parade. Flags and Shop windows were draped in black, the city shut down and the conquerors observed with scorn (Horne 2007). There were also significant alterations to the material structure of Paris during the Commune with the destruction of the Vendôme Column. The monument was erected by Napoleon in celebration of the Empire in 1810. Long reviled by the left as a symbol of imperial hubris, it took nearly a month for the Commune to prepare its toppling. Communards sealed off windows for a half mile to prevent them from shattering; laying the ground where it would fall with manure and straw to prevent it from damaging the road beneath. The people of Paris were re-engineering their environments and the information it communicated to make available new collective behaviours. If physical space functions as a form of external memory, then demolishing the column can be understood as a process of forgetting.

In the days that followed the destruction of the Commune, those that held power reconfigured the city's environments to assert their dominance once again. Goncourt recalled, "there are tricolours in every window and on every carriage" (de Goncourt et al. 1962, p. 194). Tens of thousands of suspected insurgents were captured, organized into groups based on roles such as gender or mutineer. If the Communards were not summarily executed, they were marched to different areas of the city to deal with later. Dr. Murray recalls how the Communards were taken to the Palais de l'Industrie and placed in "stalls for the horses" (Murray 1871, p. 622). For those placed before the firing squad, information was embedded into materials such as material garments to more effectively organise the killing. Agents in tricolour armbands oversaw the executions of "soldiers, army deserters, who had their tunics on inside out, with their grey cloth pockets hanging by their sides, and who seemed to be already half stripped for the firing squad" (de Goncourt et al.1962, p. 192). In their wake, the executions left a considerable number of bodies to be "metabolized" for the city to resume functioning. Various materials were employed to dispose of bones and flesh. Dr. Murray wrote, "Hundreds of those shot in cold blood had been placed in heaps and covered with earth. Chloride of lime, and also tar, had been freely used as disinfectants." (Murray 1871, p. 622) Finally, in Louise Michel's words, "Paris was an immense abattoir where, after eight days of slaughter, the hordes of flies over the mass graves put an end to the killings for they feared plague" (Michel and Mitchell 2010, p. 63). The French elite used brute force to shatter the relationships that made the Commune possible, asserting dominance over the bodies of rebellious Parisians through imprisonment and mass graves.

Marx's Bees

No human being is alive in Paris that participated in the Commune, yet their lives and actions are embedded into the material form of the city. Those past lives, like sediment, structure the environmental context of all contemporary Parisians. Generations after generations have altered the city's environments to facilitate their communal relationships. These experiences have been embedded into the epistemic landscape of Paris shaping future communities. It is in these regards the city can be seen as an act of niche construction in which the organisms modify their environments in order to shape and support their development (Laland et al. 2000). The brief yet remarkable events of the Paris Commune, bring into stark relief the entanglement of environment, social organization and cognition. In part, this is because Paris from September of 1870 to May of 1871 provides a relatively closed, clearly delineated system. During this period, human cognitive abilities, in the context of the broader transformations of the nineteenth century, intersected with a highly stressed environment to produce a unique social form. "Reproductive isolation" Manuel de Landa explains "acts as a ratchet mechanism' that conserves the accumulated adaptations and makes impossible for a given population to 'de-evolve'" (de Landa 1997, p. 61). Parisian cultural adaptations began to diverge from France as the Prussian siege fractured the "armature" of energy, matter and information supporting Second Empire Paris.

Paris' cognitive ecology was composed of a multitude of cognitive niches, from specific districts to particular institutions such as cafes, layered at different scales throughout the city. Some like the Hôtel de Ville provided larger, more deeply embedded structure, while others, like political clubs, provided smaller, more specific cognitive environments. Louise Michel gives an example from the Montmartre Vigilance Committee that shows the subtle plasticity of such environments, "I still have an old map of Paris that hung on the wall of our meeting room. I carried it back and forth across the ocean with me as a souvenir. With ink we had blotted out the empire's coat of arms, which desecrated it and which would have dirtied our headquarters" (Michel and Nic 2004, p. 51). Actions such as the tearing of the flag, tying branches and leaves on rifles, and obscuring the coat of arms were all means of modifying the environment in order to facilitate alternative possibilities. The purpose of such "epistemic actions" as Malafouris explains, "is not simply to alter the world so as to advance physically toward some goal, but rather to alter the world so as to help make available a new way of thinking about it" (Malafouris 2013, p. 194). Here we see thinking not as a disembodied abstraction, but as causal process constantly mediated through the environment. The physical configuration of the city structured flows of energy-matter in ways that bound together specific cognitive environments and elicited particular types of behaviour such as consumption or craftwork. As the Prussians altered these flows, certain bonds and continuity broke down while others emerged, such as the relationship between fuel, theatres and clubs. Parisians responded to these environmental conditions by modifying the materials of the city to communicate and facilitate new forms of organization. As Edwin Hutchins points out, "cultural things provide the meditational means to domesticate the embodied imagination" (Hutchins 2010, p. 99). Through processes of feedback, the breakdown of the urban ecology led to the transformation of the city's cognitive environments and the expansion of innovative networks of organization.

The Siege of Paris formed a bottleneck that drastically altered the way in which energy-matter flowed through Paris. With the collapse in resources necessary to sustain human bodies, many Parisians accessed the cognitive niches provided by revolutionary and socialist culture in the nineteenth century. This allowed radicalized Parisians to explore alternative organizational forms, helping sustain communities through various hardships by either meeting their basic needs, or demanding for them more effectively. These alternative social forms developed from a manipulable structure of artefacts that enabled Parisians to adapt to their environments. Things such as National Guard Cannons bound Parisians into new social relationships and provided the scaffolding with which to understand such relationships. When the National Assembly attempted to take the cannons, they challenged a densely interconnected and resilient form of organization that was operating as a functional alternative. The social conflict over these artefacts reconfigured Parisian power, catalysing the emergence of a historically significant and distinctive cognitive ecology. "All power derives from energy—it is energy put to work" notes Edmund Russell and a number of colleagues, "People gain power by enlisting people, nonhuman nature, ideas, and technology into networks supporting their goals, and energy courses through these networks" (Russell et al. 2011, p. 259). Marginalized Parisians, by organizing new ways of sustaining themselves through the siege, had tapped new sources of energy within their communities, allowing them to express new power and collective agency. Ian Hodder argues that "power is the differential flow of matter, energy and information through entanglements" (Hodder 2012, p. 214). The radicalized Parisians comprising the Commune were in a sense the successional ecology, recolonizing Paris after a disturbance, reconfiguring the environment. Ultimately, these alternative social forms were exterminated to the best of the Versaillais' ability. However, by coalescing so much socialist engagement, the Commune became its own artefact for future revolutionaries to structure their thinking around.

Ultimately, the Paris Commune helps us to understand cities as social-cognitive ecologies. Cultural modification of the environment affords the possibility of shaping developmental conditions in ways that support self-organising, collective behaviour among large populations. Urban structures sustain and facilitate the interaction of vast numbers of humans; bodies that consume resources and excrete waste, that need to regulate temperature, sleep, socialise, reproduce and so forth. Thus, the city can be seen as co-evolutionary structures that forms through human adaptation to environments generated by dense human habitation. The environmental pressures that the city places upon the individual are unique in the respect that they have taken shape to sustain the human interactions from which they emerge. To reframe this, one of the most important features of the urban environment is its ability to facilitate the adaptation of the individual to an ecology dominated by the physical-cognitive presence of its own species. As the anthropologist Tim Ingold argues, an organism is "not as a discrete, pre-specified entity but as a particular locus of growth and development within a continuous field of relationships" (Ingold 2004, p. 219). In these terms, cities are exceptionally dense cognitive ecologies that create new patterns of biological organization. Consider the walls of Paris like the mineral membrane of a cell. Within the city, matter, energy, and information intersect in a continuously generative process, flowing into various configurations of bodies, brains and materials. Eddies form in cafes and print shops, swirling on to homes, militias, and government offices. As powerful flows of energy-matter carve the channels of history, barricades emerge and great currents sweep them away, bones become sediment. By the mid-nineteenth century, Paris drew vast amounts of energy and matter into its core, a city radically reshaped by Haussmanization, the most ambitious urban renovation program of the century. The class hierarchies, and the cultural values that bound them together had consolidated with a gentrified core and solidly working class periphery. Socialist engagement provided a critical means for interpreting and engaging with the new environments of industrialization. Throughout working class districts, dedicated radicals experimented with alternative social models such as cooperatives, unions and political clubs. The so-called "ring of red," the working class fauburgs that surrounded the city, provided a fertile cognitive ecology in which the tradition, discourses and practices of the radical left could replicate and spread. While the working class periphery provided a diverse environment of cultural practices, the bourgeois core worked to maintain the stability of their system. It was a complex and nuanced ecology. Over a million soldiers, swarming out of Germany and surrounding the city, proved to be just the disturbance necessary to transform Paris. Through the months of siege and its aftermath, the relational structure of Parisian communities drastically altered. The changing flows of energy were mediated through brains, bodies, neighbourhoods, cannons, clubs, barricades, flags, food, and fuel. The Commune ignited from the materials of the city like a wildfire, radically altering the environment before its flame was extinguished. On the final day of the Commune, May 28, Edmund de Goncourt made his way to the Hotel de Ville. The building was a smouldering ruin after the fighting. The statues and clock had been shattered, the stones of the facade scorched. Amid the twisted and broken building, the nexus of Parisian revolutionary memory, stood an unscathed marble plaque inscribed with the words, “Liberté, Égalité, Fraternité.”

References

Bellinger, H., & Mitchell A. (Ed.). (2010). Men and things from the times of the commune. In *Communards: The story of the paris commune of 1871 as told by those who fought for It*. Pacifica: Erythros Press.

Blanqui, L. A. (2011). Selected works of Louise Auguste Blanqui. San Bernidino: EatDogEat Publishing.

de Boismont, B. D. (1871). A lunatic asylum during the siege of Paris. *The British Medical Journal, 1*(531).

Chakrabarty, D. (2009). The climate of history: Four theses. *Critical Inquiry, 35*(2).

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19.

Clark, A. (2008). Where brain, body and world collide. In C. Knappet & L. Malafouris (Eds.), *Material agency: Towards a non-anthropocentric approach*. New York: Springer.

Clark, A. (2011). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford: Oxford University Press.

Cluseret, G., & Mitchell A. (Ed.). (2010). Ephemera of the Siege and commune: Against the mania for braids. In*Communards: The story of the Paris commune of 1871 as told by those who fought for It*. Pacifica: Erythros Press.

Collingwood, R. G. (1951). *The idea of history*. Oxford: Oxford University Press.

de Goncourt, E., Joules de, G., & Robert, B. (Trans.). (1962). *Pages from the Goncourt journals.* New York: New York Review Books.

de Landa, M. (1997). *A thousand years of nonlinear history*. Swerve Editions. New York: Zone Books.

de Molinari, G. (1871). *Les clubs rouges pendant le siege de Paris.* Paris: Editions du Seuil.

Foucault, M. (2005). *The order of things: An archaeology of the human sciences*. London: Routledge.

Gould, R. V. (1995). *Insurgent Identities*: *Class, community, and protest in Paris from 1848 to the commune*. Chicago: University of Chicago Press.

Gulickson, G. L. (1996). *Unruly women of Paris: Images of the commune*. Ithaca: Cornell University Press.

Harison, C. (2000). The rise and decline of a revolutionary space: Paris Place de Greve and the Stonemasons of Creuse, 1750–1900. *Journal of Social History, 34*(2).

Harvey, D. (2005). *Paris: Capital of modernity*. London: Routledge.

Hodder, I. (2012). *Entangled: An archaeology of the relationship between humans and things*. Oxford: Wiley-Blackwell.

Horne, A. (2007). *The fall of Paris: The Siege and the commune 1870–71*. London: Penguin Books Ltd.

Hutchins, E. (1995). *Cognition in the wild*. Boston: MIT Press.

Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science, 2*.

Ingold, T. (2004). Beyond biology and culture. The meaning of evolution in a relational world. *Social Anthropology, 12*(2).

Johnson, M. P. (1996). *The paradise of association: Political culture and popular organization in the Paris commune of 1871*. Ann Arbor: The University of Michigan Press.

Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities, and software*. New York: Scribner.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence, 73*.

Laland, K. N., John, O.-S., & Marcus, W. F. (2000). Niche construction, biological evolution, and cultural change. *Behavioural and Brain Sciences, 23*.

Lissagaray, P.-O. (2012). *History of the Paris commune of 1871*. London: Verso.

Malafouris, L. (2013). *How things shape the mind: A theory of material engagement*. Cambridge: The MIT Press.

Marx, K., & Ben, F. (Trans.). (1990). *Capital* (Vol. I). London: Penguin Books.

Michel, L., & Nic, M. (Eds.). (2004). *Louise Michel*. New York: Ocean Press.

Michel, L., & Mitchell, A. (Ed.). (2010). Inquiry into the commune. In *communards: The story of the Paris commune of 1871 as told by those who fought for it*. Pacifica: Erythros Press.

Murray, J. (1871). Four days in the ambulances and hospitals of Paris under the commune. *The British Medical Journal, 1*(545), 622.

Reclus, E., & Mitchell A. (Ed.). (2010). The Paris commune day by day. In *Communards: The story of the Paris commune of 1871 as told by those who fought for it*. Pacifica: Erythros Press.

Renfrew, C. (2001) . Symbol before concept: Material engagement and the early development of society. In I. Hodder (Ed.), *Archaeological theory today*. Blackwell Publishing.

Russell, E., James, A., Thomas, F., & Brown, J. K. (2011). The nature of power: Synthesizing the history of technology and environmental history. *Technology and Culture, 52*(2).

Shafer, D. A. (2005). *The Paris commune: French politics, culture, and society at the crossroads of the revolutionary tradition and revolutionary socialism*. New York: Palgrave Macmillan.

Smail, D. L. (2008). *On deep history and the brain*. Berkley: University of California Press.

Stotz, K. (2010). Human nature and cognitive–developmental Niche construction. *Phenomenology and the Cognitive Sciences, 9*(4).

Varela, F. J., Rosch, E., & Thompson, E. (1992). *The embodied mind: Cognitive science and human experience*. Cambridge: The MIT Press.

Washburne, E., & Michael, H. (Ed.). (2012). *Elihu Washburne: The diary and letters of America's Minister to France during the siege and commune of Paris*. New York: Simon & Schuster.

Chapter 12
Computer-Mediated Trust in Self-interested Expert Recommendations

Jonathan Ben-Naim, Jean-François Bonnefon, Andreas Herzig, Sylvie Leblois and Emiliano Lorini

Abstract Important decisions are often based on a distributed process of information processing, from a knowledge base that is itself distributed among agents. The simplest such situation is that where a decision-maker seeks the recommendations of experts. Because experts may have vested interests in the consequences of their recommendations, decision-makers usually seek the advice of experts they trust. Trust, however, is a commodity that is usually built through repeated face time and social interaction, and thus cannot easily be built in a global world where we have immediate internet access to a vast pool of experts. In this article, we integrate findings from experimental psychology and formal tools from Artificial Intelligence to offer a preliminary roadmap for solving the problem of trust in this computer-mediated environment. We conclude the article by considering a diverse array of extended applications of such a solution.

Introduction

Important decisions are rarely made in isolation. Even when a single agent has the final say about what is to be done, the knowledge and information processing relevant to a complex decision are often distributed among several agents. Typically, one agent (the decision-maker) relies on one or several other agents (the experts) to provide recommendations based on their knowledge and know-how about the problem at hand.

J. Ben-Naim · A. Herzig
CNRS and Université de Toulouse, Toulouse, France

S. Leblois · E. Lorini
Université de Toulouse, Toulouse, France

J.-F. Bonnefon (✉)
Toulouse School of Economics, Center for Research in Management, Institute for Advanced Study in Toulouse, Centre National de la Recherche Scientifique, 21 allée de Brienne, 31000 Toulouse, France
e-mail: jean-francois.bonnefon@tse-fr.eu

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_12

The problem with experts, though, is that they may well have vested interests in the consequences of their recommendations. Think about investing. All of us who are not investment savy might want to get some expert recommendations about what to do with our savings. Sometimes, our banker is willing to provide such recommendations. We are likely to take these recommendations with a grain of salt, though, because we are aware that the banker may have vested interests in pushing some specific financial products. We are facing a dilemma between our need for expert recommendation and the potentially self-interested character of the recommendations we can get from experts, who have vested interests in the decision we are going to make from their recommendation.

Our need for expertise thus makes us the potential targets of deception from self-interested experts. The traditional solution to this dilemma is to seek the recommendations of these experts and only these experts whom we endow with our *trust*. Trust is a multidimensional concept that has been informally defined as, e.g., "the expectation that the person is both competent and reliable, and will keep your best interest in mindmind" (Barber 1983), or, quite similarly, as a combined judgment of the integrity, ability, and benevolence of an individual (Mayer et al. 1995).

Trust is a commodity that is often built through repeated social interactions (Ferrin et al. 2006; Kramer 1999). Not only do people trust other people as a function of their interpersonal history, but even the most subtle aspects of face-to-face interaction can contribute to judgments of trustworthiness. For example, people are more ready to trust interaction partners who mimic their behavioral mannerisms (Maddux et al. 2008). Whether or not this is a sensible way to endow someone with trust is, of course, a debatable question. The point is, though, that behavioral mimicry requires face-to-face interaction, and that, generally speaking, feelings of trust commonly require a history of social interaction with the person whom is to be trusted.

This solution to the problem of trust is adapted to a small world where experts on a given topic are few and personally known to the decision maker. However, in our global village, an inexhaustible pool of experts on just any given topic is always just one mouse click away from us. Whatever our concern is, the internet gives us a fast and convenient access to a vast number of experts. That would be good news, if only we knew which experts we could trust. The traditional solution to the problem of trust (repeated face time and social interaction) is no longer available in our global cognitive world.

In this article, we consider the problem of seeking expert advice through a web-based platform, where users are declared experts in various domains. We offer a list of suggestions for solving the problem of trust in this environment. The power of our approach resides in its multidisciplinarity, as we combine the cognitive insights of psychology with the formal methods of artificial intelligence to reach an integrated perspective on our problem. In the solution that we envision, regular users alternatively play the role of advisor or advisee in their interactions, depending on whom is in possession of the expert knowledge required by the situation. After each interaction, advisees have the possibility to appraise their advisor on the various dimensions that form the multifaceted concept of trust. The platform keeps a memory of these appraisals from which it can extract an aggregated, global index of the trustworthiness of any user, or decompose this global index into sub-indices corresponding to the various components of trust. Any new or regular user can thus attain a computer-mediated judgment of the extent to which any expert on the platform is to be trusted, or seek an expert whose detailed characteristics are optimally balanced to serve their needs.

In the rest of this article, we give a more detailed characterization of our problem, and we address in turn the various ingredients we need to sketch a solution. Section "Problem Specification" defines the problem of attaining a computer-mediated, complex judgment of trust within a multi-user, web-based platform of potentially self-interested experts. Section "Psychological Treatment" reviews experimental findings and psychological insights into the components of trust, their socio-cognitive antecedents, and their behavioral consequences. Section "Formal Treatment" builds on these materials and on the current state of the art in artificial intelligence to sketch a formal solution to our problem, which integrates the psychological constraints previously identified. Finally, Sect. "Extended Applications" considers various extended applications of our suggestion for computer-mediated trust.

Problem Specification

Let us imagine that you came in possession of a banjo, which you would like to sell, but whose monetary value you have no idea of. One option would be to go to the closest (and probably the only) banjo store you know of, and to ask the owner to appraise your banjo. The problem, though, is that the owner is not only the person whom you can ask about the value of your banjo, but also the person you are likely to sell the banjo to. Not knowing whether you can trust the owner not to take advantage of the situation, you turn to a web-based platform for musical instruments amateurs, where you are likely to find plenty of users who can appraise a banjo, and plenty of potential banjo buyers. Your trust problem, though, is just demultiplicated, because these are likely to be broadly the same persons. The fact that you now have an abundance of experts you might solicit is no improvement over your previous situation, because you do not have the time, the resources, or the motivation to engage in repeated social interactions with all these people in order to find out who you can trust.

We believe that the platform should offer a solution to achieve the same results as repeated social interaction. It should provide you with the basic parameters that form the building blocks of trust, as well as some index of the extent to which you can trust your potential advisors. We believe this service can be achieved by formalizing the notion of trust, and taking advantage of the history of the advisor-advisee interactions on the platform.

Not every user of the platform is an expert of everything. To continue our musical instruments exemple, some users may declare expertise in appraising banjos, whilst others may declare expertise in appraising cellos. Thus, depending on the situation, a given user may be in a position to give expert advice, or to receive it. Now consider that everytime a user x receives expert advice from another user y, x is given the opportunity to appraise this advice on all the dimensions that the complex notion of trust is known to encompass. The platform records this interaction as a tuple $R_{xy} < r_1, \ldots, r_n >$, where $r_1, \ldots, r_n$ are the appraisals given by x about the recommendation of y on the various dimensions of trust. Soon enough, the platform should be in a position to answer a request about the trustworthiness of agent y, by aggregating the information contained in the tuples expressed about y.

A number of problems must be solved to achieve such a result. First, we need to decide on the exact nature of the appraisals $r_1, \ldots, r_n$. Then, we need to decide on the way these ratings should be aggregated, both at the individual level and at the collective level. Finally, we need a formal characterization of all the components of trust and of the properties one can used to reason about them, in order to generalize our solution to environments where artificial agents interact with human agents, or among themselves. Solving these problems requires a multidisciplinary approach, drawing on experimental psychology as well as artificial intelligence methods. We now consider in turn the insights given by these two disciplines.

Psychological Treatment

Various definitions of trust can be found in the psychological literature. Some authors define trust mostly in terms of its behavioral consequences, e.g., ‘Trust is the extent to which a person is confident in, and willing to act on the basis of, the words, actions, and decisions of another’ (McAllister 1995), or trust is ‘the willingness to accept vulnerability based upon positive expectations about another’s behavior’ (Rousseau et al. 1998). Early structural perspectives on trust distinguished between trust based on cognition and trust based on affect (Johnson-George and Swap 1982; Lewis and Weigert 1985; Rempel et al. 1985). ‘Cognitive’ trust is based on explicit knowledge and ‘good reasons’ to think that a person is reliable or dependable. ‘Affective’ trust is based on an emotional bond between individuals. Clearly, just as behavioral mimicry, emotional bonds are not within the scope of our application. We should thus focus on that sort of trust which is based on explicit knowledge and deliberative thought.

Idealilly suited for our purpose is the suggestion that trustworthiness is a three-dimensional attribute composed of competence, benevolence, and integrity (Barber 1983; Mayer et al. 1995). Competence reflects the ability of a person with respect to the task at hand. Benevolence reflects a positive attitude towards the truster, and a genuine concern for the truster’s interests. Integrity reflects the adherence of the trustee to an appropriate set of ethical principles. Let us now consider in turn these three components of trust, and their potential importance in situation where advice is given.

Competence

Many studies have investigated the influence of an advisor's perceived competence on the uptake of her recommendations. Perhaps unsurprisingly, these studies concur that the recommendation of an advisor is more influential when her perceived expertise is greater. Interestingly, people seem ready to accept claims of expertise at face value, even in experimental situations where the quality of the offered 'expert' advice is actually weak (Harvey and Fischer 1997). While it is clear why people seek the advice of individuals they believe to be more competent than they are, we note that people are sometimes ready to seek the advice of individuals they believe to be *less* competent than they are; in particular, when the stakes of the decision are serious enough that they want to share the responsibility for the decision, whatever the relative expertise of their advisor (Harvey and Fischer 1997).

People appear to use a variety of cues to appraise the expertise or competence of an advisor. For example, advisors who express high confidence in their recommendation are perceived as more competent, and their recommendation is given more weight by the decision maker (Sniezek and Buckley 1995; Van Swol and Sniezek 2001). Likewise, advisors who give very precise recommendations (as opposed to vague estimates) are perceived as more competent, and, again, their recommendation is given more weight by the decision maker (Yaniv and Kleinberger 2000). All other things being equal, these strategies do appear to increase the quality of the decision making, for there seems to be an ecologically valid correlation between expertise, confidence, and precision (Sniezek and Buckley 1995; Van Swol and Sniezek 2001; Yaniv et al. 1991). Then again, these studies did not control for the possibility that the advisor has vested interests in the decision of the advisee; and a self-interested advisor may well express a very precise recommendation with great confidence, only to better serve her own interests.

Finally, a reputation for expertise is hard to build, but rapidly destroyed (Slovic 1993; Yaniv and Kleinberger 2000). Many useful recommendations are required before one is trusted as a competent advisor, but only a few average or bad recommendations are enough to lose that reputation. This phenomenon can be related to the more general *negativity bias* in impression formation (Ito and Cacioppo 2005; Skowronski and Carlston 1989). The negativity bias refers to the greater weight we attribute to negative behaviors when inferring personality traits: For example, fewer negative behaviors are needed to infer a negative trait, compared with the number of positive behaviors we need to infer a positive trait. The negativity bias, and its specific consequences for the dynamics of trust, can be conceived as a safeguard for a species that exhibits a strong tendency to spontaneous cooperation, ensuring that untrustworthy partners are quickly detected and unprofitable cooperation promptly forsaken.

Benevolence

Whenever a conflict of interest is possible, and even when it is not, people are concerned about the degree to which their advisors really care about their interests. A benevolent advisor genuinely cares about the best interests of the advisee, has a positive attitude towards the advisee, and thinks about the advisee's interests at least as much as her own.

Even when the advisor has no explicit vested interest in the situation, benevolence can contribute to trustworthiness independently of competence. For exemple, the mere fact that the advisee already knows the advisor (a proxy for benevolence) makes a difference to the advisor's perceived trustworthiness, even when controlling for the advisor's expressed confidence in her advice (a proxy for competence); in fact, this expressed confidence no longer affects trustworthiness as soon as the advisor and the advisee know each other (Van Swol and Sniezek 2001). In these experiments, an increase in trustworthiness translated into a greater weight put on the advisor's recommendation. Other experimental studies directly made it clear to decision makers whether or not some advisor was benevolent, concerned about their best interests. These experiments concured that recommendations from benevolent advisors are given greater weight in the decision (White 2005).

Interestingly, it has been claimed that people are ready to trade off competence for benevolence when the emotional load of their decision is high (White 2005). One experiment put subjects in a situation to decide whether they would leave their savings in a badly performing fund, or take them out. In the low emotional load condition, the savings were meant to pay for a summer band camp for young musicians. In that case, subjects sought competent rather than benevolent advisors. In the high emotional load condition, the savings were meant to pay for college. In that case, subjects sought benevolent advisors, and were ready to sacrifice some level of competence in order to ensure benevolence.

Whether this effect is truly due to emotional load or to another confounded variable is not quite clear, but the possibility of a trade off between competence and benevolence would already be especially relevant to our current purpose, given that we conceptualise competence and benevolence as different dimensions of the complex concept of trust. It would mean that a global index of trust might not be precise enough to accommodate people's needs. Indeed, different situations may require different mix of competence and benevolence, although the global index of trust would remain the same.

Benevolence-based trust can obviously be harmed by malevolent behavior. However, it can be repaired on the long term by subsequent benevolent behavior, or, on the short term, by promises to adopt a benevolent behavior. Apologies for malevolent behavior do not seem sufficient, though, to repair trust (Schweitzer et al. 2006).

Integrity

The integrity of the advisor reflects her unconditional adherence to a set of principles deemed appropriate by the advisee. Note that integrity so defined can be independent of benevolence. For example, one may expect an advisor to maintain confidentiality whether or not one believes the advisor to be benevolent. Conversely, one may question whether an advisor can be trusted to maintain confidentiality, independently of whether this advisor is benevolent or not.

Some indices of trust put a strong emphasis of integrity. For example, recent studies investigating the relation between emotion, trust, and the uptake of advice (Dunn and Schweitzer 2005; Gino and Schweitzer 2008) used a measure of trust that focused on whether the advisor could be expected to unconditionally honor commitments, and whether the advisor could be expected to unconditionally tell the truth. These studies found that incidental emotions (i.e., which were felt independently of the advisor) could affect this integrity-based trust, which affected in turn the weight given to the advisor's recommendation. More specifically, incidental anger decreased integrity-based trust, and incidental gratitude or happiness increased integrity-based trust.

Finally, integrity-based trust seems hard to repair once it has been harmed by a failure to honor one's commitment (Schweitzer et al. 2006). Once an individual has failed to deliver on a promise, her trustworthiness appears to be durably impaired, and not significantly repaired by apologies or renewed promises to change behavior, even when these promises are genuinely honored.

Summary

Agents faced with difficult decisions often find that they do not possess all the knowledge and expertise required to make the best possible choice. A natural solution is then to seek expert recommendation about the decision; but because experts may have vested interests in the consequences of their recommendation, they need to be trusted by the agent making the decision. Our global world offers easy access to a vast pool of experts; but it does not offer the traditional guarantees of trustworthiness that come with a history of personal interaction with all these experts.

This problem of computer-mediated trust in expert recommendations clearly falls within the scope of the distributed cognition framework proposed by Hollan and collaborators (Hollan et al. 2000). Indeed, it presents the three following characteristic features:

- Cognitive processes are distributed across the members of the social group. Not only is the final decision codependent on computations made by the decision maker and by the expert, but the trust granted to the expert is itself the result of distributed computation among the users of the platform.

- Cognitive processes involve coordination between internal and external structure. To reach an overall assessment of trustworthiness, the decision maker cannot simply inquire into the judgments made by others, but must delegate some computations to the platform and coordinate with the results of this computation.
- Processes are distributed through time in such a way that the products of earlier events transform the nature of later events. Indeed, the dynamics of trust are such that events cannot be interpreted in isolation. A display of integrity, for example, has a very different impact on trustworthiness depending on whether the expert is known to have given at least one dishonest recommendation.

Overall, the computer-mediated construction of trust is a distributed cognitive process exhibiting a complex trajectory over agents, events, and time, and requires coordination with an external computational structure. It does not result, however, in any radical conceptual rewiring of the nature of mind or trust. In that sense, we offer a 'weak' distributed perspective, focused on the multi-level aggregation of the cognitive outputs of humans and artefacts: a formally difficult problem, but a tractable one.

Our approach sticks to a conceptualisation of the mind as an information processing system, with clearly defined inputs and outputs; and our work rests on the assumption that a significant portion (though clearly not the whole) of trust-building boils down to information processing. Although some aspects of trust-building elude our formalization, we believe that the cold information processes captured by our formalization can already offer some solid decisional grounds. These processes are constrained by a number of variables and psychological dimensions, which we explored in the previous section. In line with previous psychological research, we conceptualise trust as a multidimensional concept comprising competence (expert ability), benevolence (positive attitude and concern towards the interests of the advisee), and integrity (unconditional adherence to a set of principles deemed appropriate by the advisee).

These three dimensions of trust exhibit different degrees of asymmetry in the differential impact of positive and negative information. In the case of integrity, negative information receives extremely greater weight than positive information. This asymmetry is also observed with respect to competence, but apparently to a lesser extent. Finally, the asymmetry would appear to be the least pronounced in the case of benevolence.

Some compensation seems possible between the dimensions of competence and benevolence, since situations appear to exist where advisees are willing to sacrifice some measure of competence to ensure some mesure of benevolence. It is less clear whether integrity can be traded that way, or whether it should be considered as a completely non-compensatory dimension of trust. One possibility, that would need empirical validation, is that the level of integrity functions as the upper-bound for the level of trustworthiness. A related solution to the problem of computer-mediated trust is to first filter out advisors who have been judged to lack integrity; and then to provide the user with aggregated indices of competence and benevolence, without taking the responsibility to trade one for another in a global index of trustworthiness. This responsibility should be left to the user, who knows best whether the situation primarily calls for competence, benevolence, or both.

We now turn to the formal treatment of our problem. We introduce a logical framework wherein the three aspects of trust can be formally characterized, and wherein we can model trust reasoning about these three aspects.

Formal Treatment

This section presents a logical framework called $\mathcal{TRUST}$ in which the competence, benevolence and integrity of an advisor can be formally characterized. $\mathcal{TRUST}$ is a multi-modal logic which supports reasoning about time, agents' actions and agents' mental attitudes including beliefs and goals. It also allow one to express the normative concept of obligation. In this sense, $\mathcal{TRUST}$ combines the expressiveness of dynamic logic (Harel et al. 2000), temporal logic (Emerson 1990) and deontic logic (Åqvist 2002) with the expressiveness of a so-called BDI (belief, desire, intention) logic of agents' mental attitudes (Cohen and Levesque 1990). We introduced the logic $\mathcal{TRUST}$ in our previous works on the logical formalization of the concepts of trust and reputation (Lorini and Demolombe 2008). It is not the aim of this work to discuss the precise semantics of the modal operators of the logic $\mathcal{TRUST}$. We just present them in an informal way by highlighting their intuitive meanings and their basic properties.[1]

The syntactic primitives of the logic $\mathcal{TRUST}$ are the following:

- a nonempty finite set of agents $AGT = \{i, j, \ldots\}$;
- a nonempty finite set of atomic actions $AT = \{a, b, \ldots\}$;
- a finite set of propositional atoms $ATM = \{p, q, \ldots\}$.

The language of $\mathcal{TRUST}$ is defined as the smallest superset of ATM such that:

- if $\varphi, \psi \in \mathcal{L}$, $\alpha \in ACT$ and $i \in AGT$ then $\neg\varphi, \varphi \vee \psi, \mathtt{Does}_{i:\alpha}\varphi, \mathtt{Bel}_i\varphi, \mathtt{Choice}_i\varphi, \mathtt{Past}\varphi, \mathtt{Obl}\varphi \in \mathcal{L}$.

ACT is the set of complex actions and is defined as follows:

$$ACT = AT \cup \left\{ inf_j(\varphi) | j \in AGT,\ \varphi \in \mathcal{L} \right\}.$$

An action of the form $inf_j(\varphi)$ denotes the action of informing agent j that φ is true. We call this kind of actions informative actions.

---

[1] See for instance (Lorini and Demolombe 2008) for an analysis of the semantics of these operators, their relationships, and their correspondence with the structural conditions on the models of the logic $\mathcal{TRUST}$

Thus, the logic $\mathcal{TRUST}$ has five types of modalities: $\mathtt{Bel}_i$, $\mathtt{Choice}_i$, $\mathtt{Does}_{i:\alpha}$, $\mathtt{Past}\varphi$ and $\mathtt{Obl}$. These modalities have the following intuitive meaning.

- $\mathtt{Bel}_i\,\varphi$: the agent *i* believes that φ;
- $\mathtt{Does}_{i:\alpha}\,\varphi$: agent *i* is going to do α and φ will be true afterward ($\mathtt{Does}_{i:\alpha}\mathrm{T}$ is read: agent *i* is going to do α);
- $\mathtt{Past}\varphi$: it has at some time been the case that φ;
- $\mathtt{Choice}_i\,\varphi$: agent *i* has the chosen goal that φ holds (or simply agent *i* wants that φ holds).

Operators of the form $\mathtt{Choice}_i$ are used to denote an agent's chosen goals, that is, the goals that the agent has decided to pursue. We do not consider how an agent's chosen goals originate through deliberation from more primitive motivational attitudes called desires (see e.g. (Castelfranchi and Paglieri 2007; Conte and Castelfranchi 1995; Rao and Georgeff 1991) on this issue).

The following abbreviations will be convenient:

$$\mathtt{Intends}_i(\alpha) \overset{def}{=} \mathtt{Choice}_i\mathtt{Does}_{i:\alpha}\mathrm{T}$$

$$\mathtt{Inf}_{i,j}(\varphi) \overset{def}{=} \mathtt{Does}_{i:inf_j(\varphi)}\mathrm{T}$$

$$\mathtt{BelIf}_{i\varphi} \overset{def}{=} \mathtt{Bel}_{i\varphi} \vee \mathtt{Bel}_i\neg\varphi$$

$\mathtt{Intends}_i(\alpha)$ stands for 'agent *i* intends to do action α'. This means that *i*'s intention to perform action α is defined by agent *i*'s choice to perform action $\alpha.\mathtt{Inf}_{i,j}(\varphi)$ stands for 'agent *i* informs agent *j* that the fact φ is true'. Finally, $\mathtt{BelIf}_i\varphi$ stands for 'agent *i* believes whether φ is true'.

Operators for actions of type $\mathtt{Does}_{i:\alpha}$ are normal modal operators satisfying the axioms and rules of inference of the basic normal modal logic K (Chellas 1980).

Operators of type $\mathtt{Bel}_i\,\varphi$ are just standard doxastic operators in Hintikka style (Hintikka 1962) satisfying the axioms and rules of inference of the so-called system KD45 (Chellas 1980). It follows that an agent cannot have inconsistent beliefs, and an agent has positive and negative introspection over his beliefs. Formally:

$\mathbf{D}_{Bel}$ $\neg(\mathtt{Bel}_i\,\varphi \wedge \mathtt{Bel}_i\neg\varphi)$
$\mathbf{4}_{Bel}$ $\mathtt{Bel}_i\,\varphi \rightarrow \mathtt{Bel}_i\mathtt{Bel}_i\,\varphi$
$\mathbf{5}_{Bel}$ $\neg\mathtt{Bel}_i\,\varphi \rightarrow \mathtt{Bel}_i\neg\mathtt{Bel}_i\,\varphi$

As emphasized above, operators of the form $\mathtt{Choice}_i$ express an agent's chosen goals. These are similar to the modal operators studied in (Cohen and Levesque 1990). Since an agent's chosen goals result from the agent's deliberation, they must satisfy two fundamental rationality principles: chosen goals have to be consistent (i.e., a rational agent cannot decide to pursue inconsistent state of affairs); chosen goals have to be compatible with the agent's beliefs (i.e., a rational agent cannot decide to pursue something that it believes to be impossible). Thus, every operator $\mathtt{Choice}_i$ is supposed to satisfy the axioms and rules of inference of the so-called system KD (Chellas 1980). It follows that an agent cannot choose φ and $\neg\varphi$ at the same time. Moreover chosen goals have to be compatible with beliefs. Formally:

D$_{Choice}$ $\neg(\texttt{Choice}_i\varphi \wedge \texttt{Choice}_i\neg\varphi)$
Comp$_{Bel,Choice}$ $\mathrm{Bel}_i\varphi \rightarrow \neg\mathrm{Choice}_i\neg\varphi$

As far as the modal operator for obligation is concerned, we take the operator of Standard Deontic Logic (SDL) (Åqvist 2002). That is, the modality `Obl` is also supposed to satisfy the axioms and rules of inference of the so-called system KD. It follows that obligations have to be consistent. That is:

D$_{Obl}$ $\neg(\texttt{Obl}\varphi \wedge \texttt{Obl}\neg\varphi)$

The temporal operator $\texttt{Past}\varphi$ is also a normal modality which satisfies the axioms and rules of inference of system of the basic normal modal logic K. The following two additional axioms are added in order to capture two essential aspects of time.

4$_{Past}$ $\texttt{PastPast}\varphi \rightarrow \texttt{Past}\varphi$
Connected$_{Past}$ $(\texttt{Past}\varphi \wedge \texttt{Past}\psi) \rightarrow (\texttt{Past}(\varphi \wedge \texttt{Past}\psi) \vee \texttt{Past}(\psi \wedge \texttt{Past}\varphi) \vee$
$\texttt{Past}(\psi \wedge \varphi))$

Axiom 4_{Past} says that the past satisfies transitivity: if it has been the case that it has been the case that φ then it has been the case that φ. Axiom **Connected** $_{Past}$ just expresses that the past is connected: if there are two past moments t and t' then either t is in the past of t' or t' is in the past t or $t = t'$.

Other relationships between the different modalities of the logic $\mathcal{TRUST}$ are expressed by the following logical axioms.

Alt$_{Does}$ $\texttt{Does}_{i:\alpha}\varphi \rightarrow \neg\texttt{Does}_{j:\beta}\neg\varphi$
IntAct $\texttt{Does}_{i:\alpha}\top \rightarrow \texttt{Intends}_i(\alpha)$
Inc$_{Time,Does}$ $(\varphi \wedge \texttt{Does}_{i:\alpha}\top) \rightarrow \texttt{Does}_{i:\alpha}\texttt{Past}\varphi$

Axiom **Alt**$_{Does}$ says that: if agent i is going to do α and φ will be true afterward, then it cannot be the case that agent j is going to do β and $\neg\varphi$ will be true afterward. Axiom **IntAct** relates an agent's intentions with his actions. According to this axiom, an agent is going to do action α only if has the intention to perform action α. In this sense it is supposed that an agent's *doing* is by definition intentional. A similar axiom has been studied in (Lorini and Herzig 2008; Lorini et al. 2006) in which a logical model of the relationships between intention and action performance is proposed. Finally Axiom **Inc**$_{Time,Does}$ expresses that every action occurrence goes from the present to the future (i.e. actions do not go back to the past). That is, if φ is true in the present and agent i does action α then, after the occurrence of action α, φ is true at some point in the past.

Formal Definitions of Competence, Benevolence and Integrity

The aim of this section is to formalize in the logic $\mathcal{TRUST}$ the three properties competence, benevolence and integrity of a potential advisor.

We start with the concept of competence of an advisor to provide good recommendations about a certain issue φ.

Definition 1 ***Competence*** *Agent j is a competent advisor (or competent information source) about a certain issue φ if and only if, if agent j believes that φ then φ is true.*

This notion of competence can be formally expressed as follows:

$$\mathtt{Competent}_j(\varphi) \overset{def}{=} \mathtt{Bel}_j\,\varphi \rightarrow \varphi.$$

The second concept we aim at formalizing is benevolence.

Definition 2 ***Benevolence*** *Agent j is a benevolent advisor (or benevolent information source) about a certain issue φ if and only if, for every agent i, if j believes that i wants to believe whether φ is true and j believes that φ is true then j informs i about his opinion.*

This notion of benevolence can be formally expressed as follows:

$$\mathtt{Benevolent}_j(\varphi) \overset{def}{=} \bigwedge_{i\in AGT} \big((\mathtt{Bel}_j\mathtt{Choice}_i\mathtt{BelIf}_i\varphi \wedge \mathtt{Bel}_j\,\varphi) \rightarrow \mathtt{Inf}_{j,i}(\varphi)\big).$$

As far as integrity is concerned, we split this concept into three different concepts of sincerity, confidentiality and obedience. That is, we suppose that the expression ‘the advisor satisfies the property of integrity’ means that the advisor is sincere, obedient, and he guarantees the confidentiality of the information.

Definition 3 ***Sincerity*** *Agent j is a sincere advisor (or sincere information source) about a certain issue φ if and only if, for every agent i, if j informs i that φ is true then j believes that φ is true.*

This notion of sincerity can be formally expressed as follows:

$$\mathtt{Sincere}_j(\varphi) \overset{def}{=} \bigwedge_{i\in AGT} \big(\mathtt{Inf}_{j,i}(\varphi) \rightarrow \mathtt{Bel}_j\,\varphi\big).$$

Definition 4 ***Confidentiality (or Privacy)*** *Agent j is an advisor (or information source) which guarantees the confidentiality (or privacy) of the information φ if and only if, for every agent i, if it is obligatory that j does not inform i that φ is true then j does not inform i that φ is true.*

This notion of confidentiality can be formally expressed as follows:

$$\mathtt{Privacy}_j(\varphi) \overset{def}{=} \bigwedge_{i\in AGT} \big(\mathtt{Obl}\neg\mathtt{Inf}_{j,i}(\varphi) \rightarrow \neg\mathtt{Inf}_{j,i}(\varphi)\big).$$

Definition 5 ***Obedience*** *Agent j is an obedient advisor (or obedient information source) about a certain issue φ if and only if, for every agent i, if j is obliged to inform i about φ then j informs i about φ.*

This notion of obedience can be formally expressed as follows:

$$\mathtt{Obedient}_j(\varphi) \overset{def}{=} \bigwedge_{i \in AGT} \big(\mathtt{Obl}\ \mathtt{Inf}_{j,i}(\varphi) \rightarrow \mathtt{Inf}_{j,i}(\varphi)\big).$$

We define the integrity of the advisor j about a certain issue φ as the logical conjunction of j's sincerity about φ, j's obedience about φ, the fact that j guarantees the confidentiality of the information φ:

$$\mathtt{Integrity}_j(\varphi) \overset{def}{=} \mathtt{Sincere}_j(\varphi) \wedge \mathtt{Privacy}_j(\varphi) \wedge \mathtt{Obedient}_j(\varphi).$$

Trust Reasoning About Competence, Benevolence and Integrity

When assessing the trustworthiness of a certain advisor k, the truster i evaluates whether k has the three properties of competence, benevolence and integrity. In many situations, such an evaluation might depend on what agent i has heard about the advisor k in the past. In particular, agent i's evaluation of an agent k's competence, benevolence and integrity might be based on what the other agents told to i about k. In these situations, agent i has to apply certain procedures for *aggregating* all information that he has received from the other agents about k's properties.

The logic $\mathcal{TRUST}$ allows to formalize some of these procedures, namely *majority* and *unanimity*. For instance, we can specify the concept of 'the majority of agents informed agent i that agent k is benevolent about φ'.

$$\mathtt{Maj}_i(\mathtt{Benevolent}_k(\varphi)) \overset{def}{=} \bigvee_{J \subseteq AGT, |J| > |AGT \setminus J|} \left(\bigwedge_{j \in J} \mathtt{Past}\ \mathtt{Inf}_{j,i}(\mathtt{Benevolent}_k(\varphi)) \right)$$

According to this definition, the majority of agents informed agent i that agent k is benevolent about φ (noted $\mathtt{Maj}_i(\mathtt{Benevolent}_k(\varphi))$) if and only if there exists a group of agents J such that every agent j in J informed i that k is benevolent about φ and J is larger than its complement with respect to AGT.

In a similar way we can express that 'the majority of agents informed agent i that agent k is competent about φ'.

$$\mathtt{Maj}_i(\mathtt{Competent}_k(\varphi)) \overset{def}{=} \bigvee_{J \subseteq AGT, |J| > |AGT \setminus J|} \left(\bigwedge_{j \in J} \mathtt{Past}\ \mathtt{Inf}_{j,i}(\mathtt{Competent}_k(\varphi)) \right)$$

As far as unanimity is concerned, we can specify the concept of ‘all agents unanimously informed agent *i* that agent *k* satisfies the property of integrity’.

$$\mathtt{Unan}_i(\mathtt{Integrity}_k(\varphi)) \overset{def}{=} \bigwedge_{j \in AGT} \mathtt{Past\ Inf}_{j,i}(\mathtt{Integrity}_k(\varphi))$$

The previous definitions of majority-based benevolence and competence and unanimity-based integrity can be used to specify the procedures adopted by agent *i* to evaluate a certain advisor *k*. From the experimental literature that we reviewed in Sect. “Psychological Treatment”, it seems sensible to use a strong unanimity procedure for integrity, but to allow a more lenient majority procedure for competence and benevolence:

$$\mathtt{Maj}_i(\mathtt{Competent}_k(\varphi)) \rightarrow \mathtt{Bel}_i\mathtt{Competent}_k(\varphi).$$

This rule says that if the majority of agents informed *i* that *k* is a competent advisor then *i* believes so.

$$\mathtt{Maj}_i(\mathtt{Benevolent}_k(\varphi)) \rightarrow \mathtt{Bel}_i\mathtt{Benevolent}_k(\varphi).$$

This rule says that if the majority of agents informed *i* that *k* is a benevolent advisor then *i* believes so.

$$\mathtt{Unan}_i(\mathtt{Integrity}_k(\varphi)) \rightarrow \mathtt{Bel}_i\mathtt{Integrity}_k(\varphi).$$

This rule just says that if all agents informed *i* that *k* is an advisor which satisfies the property of integrity then *i* believes so.

At this point, and although much has still to be articulated, we will conclude the formal analysis of our problem. Indeed, our goal in this article has not been to solve the problem of computer-mediated trust in partial expert recommendations, but rather to provide a roadmap for addressing this problem, by integrating findings from experimental psychology and formal tools from Artificial Intelligence. In the last section of this article, we go beyond our initial problem by suggesting extended applications of our approach, to a range of problems where trust (or reputation) cannot be assessed by personal interaction, where agents cannot be vouched for by an objective arbiter, but where the possibility remains of applying some variant of our approach.

Extended Applications

In its most general formulation, the problem we have tackled here concerns multi-agent applications where users have to evaluate (or simply compare) agents, but it is impossible to call on an objective arbiter to provide some help. This may happen for various reasons, for example, the number of agents is too large, no arbiter is considered sufficiently competent and sincere, arbiters are too expensive, etc. However, in such applications, a lot of feedback may be available, that is, information about agents provided by other agents. Trust and reputation systems of the kind we have envisioned here are conceived to exploit such information in order to help users to take decisions about other agents.

The information provided by peers should be used with caution. It can be incomplete, and it may be downright false. Indeed, agents may have vested interests in their judgments, and therefore may lie or hide the truth to serve their interests. Another issue that is critical in any trust and reputation system is that of *cycles* of information (e.g., *a* provides information about *b*, *b* provides information *c*, and *c* provides information *a*). Trust and reputation systems have to give different weights to the pieces of information provided by the agents, but assigning such weights in a rational way turns out to be difficult in the presence of information cycles. In this final section, we consider several situations where a trust and reputation system can be used to overcome the absence of a neutral, objective arbiter.

Currently, the best-known examples of a virtual community of agents are social networks such as of Facebook or MySpace. We briefly evoke this setting because of its popularity, although it does not, strictly speaking, relate to our topic; indeed, the reputation system that can be implemented in this setting is likely to be gratuitous (it is not meant as a decision help) and unrelated to our central issue of trust. Still, in such a social network, a wealth of information is given by agents about other agents. For example, in addition to the comments and pictures they leave on each other 'walls', users can rate their virtual friends on a number of dimensions (are they attractive? honest? serious?), or vote for the nicest person in their network; and all this information can be used to extract aggregated judgments about any particular user.

Other applications are, to a greater extent, geared to help decisions. For example, e-commerce applications like Ebay are such systems where it is useful to have information about sellers before deciding to buy an item. Here, the agents are the users and the dimension of trust that is the most decisive is integrity, the expectation that the seller respects his commitments and tell the truth. In this kind of system, there are too many buyers and sellers for an external arbiter to evaluate them all. However, after each transaction, buyer and seller have an opportunity to appraise each other. This rich amount of feedback can be exploited to reach aggregated evaluation of individual ebayers. Ebay is already equipped with a simple reputation system, which does not however explicitly measure a score of integrity-based trust. Rather, it uses a simple scheme where a positive feedback from a buyer brings one point, a negative feedback removes one point, and a neutral feedback has no consequences. Symbolic trinkets are attached to some scores (e.g., a star when the seller reaches a net score of 10 points). One limitation of this system is that it does not weight feedback according to the reputation of the ebayer who provided it.

Agents in a trust and reputation system need not be human. Indeed, web pages may be seen as agents, and a link between two pages may be construed as a positive recommendation by the linking page about the linked page. Pages that gathered the most aggregated support can be considered as more trustworthy along the competence dimension of trust: they are the pages were relevant information is to be found. This is in fact one of the broad principles that PageRank (the reputation system used by the Google search engine) is based on.

Scientific citation indices offer another application of trust and reputation systems, where scientific papers are the agents (or, perhaps, the minions of the scientists who wrote them). In most scientific reputation indices, citing a paper is construed as a positive recommendation of that paper. This is true of very basic indices such as the raw number of citations, as well as of more elaborated indices such as the *h* index. As often, this framework gives every citation the same weight in the aggregated evaluation of the paper or the scientist who wrote the collection of papers under consideration. A trust and reputation system would allow to weight a citation according to aggregated scores of the citing paper that would take into account the potential for vested interests in citing one article rather than another.

The last application we consider in this discussion is less publicized, partly because of its more technical nature. It concerns the important issue of message encryption and public key certificates. Without engaging in too technical a discussion, we can summarize the problem as follows. Various agents wish to exchange encrypted messages. Each agent is in possession of two personal keys. One of these is public, it can be used to encrypt messages sent to this agent; the other is private, it is used by the agent to decrypt messages that were encrypted using his or her public key.

One concern within this framework is that a malicious agent may assume the identity of another agent a, and pretend that her own public key is actually the public key of a. Other agents may then mistakenly believe they are encrypting messages with the public key of a, when they are really using the public key of the malicious agent. The malicious agent can then intercept and decrypt messages that were meant for a.

The problem for any agent, then, is to have sufficient ground to believe that what is advertized as the public key of a truly is the public key of a. Public key certificates are used to solve that problem. A public key certificate is a document supposedly written by an agent b, signed with a public key K_b, that certifies that an agent a is the owner of a public key K_a. Consequently, we can extract from this framework a set of pairs composed of an agent and a public key supposedly belonging to it. In addition, we can extract a binary support relation between these pairs. More precisely, we consider that a pair (b, K_b) supports the credibility of a pair (a, K_a) if there exists a certificate from b, signed with K_b, stating that a is the owner of K_a. This information can be used to evaluated the credibility of the different pairs. For example, the well-known Pretty Good Privacy system looks for chains of pairs, where the credibility of the first element is trusted by the sender, each element supports the next one, and the last element is the receiver.

The main limitation of this framework is its extreme cautiousness. If the chain of certification does not go back to some agent trusted by the potential sender, no encrypted message can be sent. One way to overcome this extreme cautiousness (at some risk), is to use the kind of trust and reputation system that we have considered through this article, and to appraise the trustworthiness of an agent-key pair based on the structure of the certification graph.

We do not consider this application in any greater detail, for our goal in this last section was rather to give a broad perspective of the various problems that can be tackled by the general approach we have outlined in this paper. We hope that the reader will have gained a sense of the many domains where a trust and reputation system can help appraise the characteristics of some agents who cannot be evaluated by a central, neutral authority. These applications must be supported by a mix of psychological findings and artificial intelligence formalisms, whose exact composition depends on the extent to which the agents in the system are human or human-like in their behavior and intentions.

References

Åqvist, L. (2002). Deontic logic. In D. M. Gabbay & F. Geunther (Eds.), *Handbook of philosophical logic*. Dordrecht: Kluwer Academic Publishers.

Barber, B. (1983). *The logic and limits of trust*. NJ: Rutgers Univ. Press.

Castelfranchi, C., & Paglieri, F. (2007). The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intentions. *Synthese, 155,* 237–263.

Chellas, B. F. (1980). *Modal logic: An introduction*. Cambridge: Cambridge University Press.

Cohen, P. R., & Levesque, H. J. (1990). Intention is choice with commitment. *Artificial Intelligence, 42,* 213–261.

Conte, R., & Castelfranchi, C. (1995). *Cognitive and social action*. London: London University College of London Press.

Dunn, J. R., & Schweitzer, M. E. (2005). Feeling and believing: the influence of emotion on trust. *Journal of Personality and Social Psychology, 88,* 736–748.

Emerson, E. A (1990). Temporal and modal logic. In J. van Leeuwen (Ed.), *Handbook of Theoretical Computer Science, Volume B: Formal Models and Sematics*. North-Holland Pub. Co./MIT Press.

Ferrin, D. L., Dirks, K. T., & Shah, P. P. (2006). Direct and indirect effects of third-party relationships on interpersonal trust. *Journal of Applied Psychology, 91,* 870–883.

Gino, F., & Schweitzer, M. E. (2008). Blinded by anger or feeling the love: How emotions influence advice taking. *Journal of Applied Psychology, 93,* 1165–1173.

Harel, D., Kozen, D., & Tiuryn, J. (2000). *Dynamic logic*. Cambridge: MIT Press.

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes, 70,* 117–133.

Hintikka, J. (1962). *Knowledge and belief.* New York: Cornell University Press.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction, 7,* 174–196.

Ito, T. A., & Cacioppo, J. T. (2005). Variations on a human universal: Individual differences in positivity offset and negativity bias. *Cognition and Emotion, 19,* 1–26.

Johnson-George, C. E., & Swap, W. C. (1982). Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology, 43,* 1306–1317.

Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology, 50,* 569–598.

Lewis, J. D., & Weigert, A. (1985). Trust as social reality. *Social Forces, 63,* 967–985.

Lorini, E. A., & Demolombe, R (2008). Trust and norms in the context of computer security. In Springer-Verlag (Ed.), *Proc. of the Ninth International Conference on Deontic Logic in Computer Science (DEON'08)*, volume 5076 of *LNCS* (pp. 50–64).

Lorini, E., & Herzig, A (2008). A logic of intention and attempt. *Synthese*, *163*(1), 45–77.

Lorini, E., Herzig, A., & Castelfranchi, C (2006). Introducing "attempt" in a modal logic of intentional action. In *Logics in Artificial Intelligence: 10th European Conference (JELIA 2006)*, volume 4160 of *LNAI* (pp. 280–292). Springer.

Maddux, W. W., Mullen, E., & Galinsky, A. D. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negociation outcomes. *Journal of Experimental Social Psychology, 44,* 461–468.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20,* 709–734.

McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal, 38,* 24–59.

Rao, A. S., & Georgeff, M. P (1991). Modelling rational agents within a BDI-architecture. In M. Kaufmann (Ed.), *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91)* (pp. 473–484).

Rempel, J. K., Holmes, J. G., & Zanna, M. D. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49,* 95–112.

Rousseau, M., Sitkin, S., Burt, R., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23,* 393–404.

Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes, 101,* 1–19.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105,* 131–142.

Slovic, P. (1993). Perceived risk, trust and democracy. *Risk Analysis, 13,* 675–685.

Sniezek, J. A., & Buckley, B. (1995). Cueing and cognitive conflict in judge-advisor decision making. *Organizational Behavior and Human Decision Processes, 62,* 159–174.

Van Swol, L. M., & Sniezek, J. A. (2001). Trust, confidence and expertise in a judge-advisor system. *Organizational Behavior and Human Decision Processes, 84,* 288–307.

White, T. B. (2005). Consumer trust and advice acceptance: The moderating roles of benevolence, expertise, and negative emotions. *Journal of Consumer Psychology, 15,* 141–148.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes, 83,* 260–281.

Yaniv, I., Yates, J. F., & Smith, J. E. K. (1991). Measures of discrimination skill in probabilistic judgment. *Psychological Bulletin, 110,* 611–617.

Chapter 13
The Cognitive Ecology of the Internet

Paul Smart, Richard Heersmink and Robert W. Clowes

Abstract In this chapter, we analyze the relationships between the Internet and its users in terms of situated cognition theory. We first argue that the Internet is a new kind of cognitive ecology, providing almost constant access to a vast amount of digital information that is increasingly more integrated into our cognitive routines. We then briefly introduce situated cognition theory and its species of embedded, embodied, extended, distributed and collective cognition. Having thus set the stage, we begin by taking an embedded cognition view and analyze how the Internet aids certain cognitive tasks. After that, we conceptualize how the Internet enables new kinds of embodied interaction, extends certain aspects of our embodiment, and examine how wearable technologies that monitor physiological, behavioral and contextual states transform the embodied self. On the basis of the degree of cognitive integration between a user and Internet resource, we then look at how and when the Internet extends our cognitive processes. We end this chapter with a discussion of distributed and collective cognition as facilitated by the Internet.

Introduction

The Internet has radically altered the way we access information, deeply transforming the way we think, act and remember. Very few of our cognitive and epistemic endeavors, either individual or collective, are undertaken without some

---

P. Smart (✉)
Electronics and Computer Science, University of Southampton,
SO17 1BJ Southampton, UK
e-mail: ps02v@ecs.soton.ac.uk

R. Heersmink
Department of Philosophy, Macquarie University, Sydney, Australia
e-mail: richard.heersmink@gmail.com

R.W. Clowes
Instituto de Filosofia da Nova, Faculdade de Ciências Sociais e Humanas,
Universidade Nova de Lisboa, Av. de Berna, 26-40 piso, 1069-061 Lisbon, Portugal
e-mail: robert.clowes@gmail.com

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_13

sort of involvement of the Internet: we look up information with search engines, store documents in the cloud, navigate with online maps, read online newspapers and books, engage with online courses, use online recipes, check online timetables, watch online videos, play online games, and so on. In this chapter we conceptualize the Internet as part of our cognitive ecology (see Bateson 1972; Hutchins 2010). Cognitive ecologies can be defined as “the multidimensional contexts in which we remember, feel, think, sense, communicate, imagine, and act, often collaboratively, on the fly, and in rich ongoing interaction with our environments” (Tribble and Sutton 2011, p. 94). In our contemporary digital society (see Lupton 2015), the Internet constitutes an important part of our cognitive ecology, as most people spend many hours per day interacting with the Internet via TVs, desktop computers, laptops, tablets, smartphones, smartwatches, and other devices.

We agree with Hutchins that “the study of cognitive ecosystems will become an increasingly important part of cognitive science” (Hutchins 2010, p. 705). This chapter therefore aims to understand the cognitive and epistemic implications of the Internet. It aims to take seriously the Internet as an important part of our cognitive ecology, contributing to human cognition in a variety of ways. Our analysis takes the perspective of situated cognition theory and its species of embedded, embodied, extended, distributed and collective cognition (Robbins and Aydede 2009). Situated cognition theory is a set of approaches to human cognition underlining the importance of our embodied interactions with the socio-technological environment (Hutchins 1995; Clark 1997, 2008; Dourish 2001; Anderson 2003; Gallagher 2005; Menary 2010; Sutton 2010; Malafouris 2013). A situated approach allows us to look at the relationship between the Internet and its users from a variety of viewpoints. It is partly for this reason that we believe this approach is in an especially strong position to help us understand the cognitive contribution of the Internet, as we want to cast our explanatory net as widely as possible. A number of recent works have, likewise, adopted a situated approach when discussing the relation between human cognition and the Internet or World Wide Web.[1] These include Clowes (2015), Halpin (2013), Halpin et al. (2010), Smart (2012, 2014), and Staley (2014). This chapter builds on and synthesizes these works into an extensive overview and analysis of the cognitive implications of the Internet.

The chapter has the following structure. We start by conceptualizing the Internet as a new kind of cognitive ecology, looking specifically at the Social Web, the Web of Data, mobile computing, human computation, augmented reality, and personal informatics (see Sect. The Internet: A New Kind of Cognitive Ecology). After that, we outline situated cognition theory, briefly analyzing some of the relations between embedded, embodied, extended, distributed, and collective cognition

---

[1] The current chapter uses the term ‘Internet’ as a catch-all term for all the various applications that are built on top of the Internet. This includes the World Wide Web, which is currently the most popular Internet application. As such, when we refer to the Internet as a cognitive ecology, we mean to suggest that the Web (as well as all other Internet applications, such as email) should be included as part of the cognitive ecology.

(see Sect. Situated Cognition Theory). We then analyze the Internet in terms of these frameworks. Taking an embedded cognition view, we conceptualize how the Internet shapes certain kinds of cognitive tasks (see Sect. Embedded Cognition). We then look at how the Internet enables new kinds of embodied interaction, extends certain aspects of our embodiment, and (potentially) transforms the embodied self (see Sect. Embodied Cognition). Thereafter, on the basis of the degree of cognitive integration between the user and online information, we analyze the kinds of situations in which the Internet might be said to form a constitutive part of human cognitive processes—part of the physical machinery of the human mind (see Sect. Extended Cognition). We end this chapter with a discussion of distributed and collective cognition as facilitated by large-scale forms of technology-mediated social participation (see Sect. Collective Cognition).

The Internet: A New Kind of Cognitive Ecology

Over the past several decades, the Internet has emerged as an important part of the material environment in which human (and perhaps machine[2]) intelligence is realized. This view of the Internet ties in nicely with ecological approaches to cognition (see Bateson 1972; Malafouris 2013; Hutchins 2010), which emphasize the role of cognitive ecosystems (i.e., complex networks of material forces and factors that span brain, body and world) in attempts to explain and understand human cognitive capabilities. Hutchins (2010), for example, suggests that our attempt to understand "cognitive phenomena must include a consideration of the environments in which cognitive processes develop and operate" (p. 706). From an ecological perspective, therefore, the Internet can be seen as part of the ecosystem for human cognition: it serves as an increasingly important part of the material environment in which an ever-expanding array of human cognitive and epistemic activities unfold. An ecological perspective also invites something of a methodological and conceptual shift when it comes to analyzing the cognitive effects of the Internet (see Smart 2013). In particular, by adopting an ecological perspective, it seems appropriate to analyze the Internet from the perspective of approaches that are typically subsumed under the heading of a situated approach to cognition, e.g., embedded, embodied, extended and distributed/collective approaches to cognition (see Robbins and Aydede 2009). Such an approach seems particularly appropriate once we consider some of the areas of research and development that are shaping the technological trajectory of the Internet and Web. Prominent areas of research

---

[2] Smart (2013), for example, suggests that the Web provides a new kind of ecological context in which advanced forms of machine intelligence might emerge.

attention in the Web and Internet Science (WAIS) community thus include (but are not limited to) the following:[3]

- **The Social Web:** Some of the most popular systems on the Web today form part of what is known as the Social Web. These include social networking sites, such as Facebook; microblogging services, such as Twitter; and social media systems, such as Wikipedia and YouTube. The Social Web is currently a major focus of interest for those concerned with issues of collective intelligence and distributed cognition (Hendler and Berners-Lee 2010; Malone et al. 2010; Chi 2009; Chi et al. 2008; Halpin 2013).
- **The Web of Data:** The Web is increasingly viewed as a platform that supports the implementation of data-driven apps, services and data analytic capabilities. The use of linked data formats promises to increase the accessibility of online content, as well as enhance the flexibility of digital representations. Such capabilities are sometimes seen as relevant to the emergence of Web-based forms of cognitive extension (Smart, in press, 2012).[4] It is important not to underestimate the transformative impact of the sort of ubiquitous data environment that might be brought into being on the back of the Web of Data. In order to help us see this, consider the way in which the constant stream of data provided by the Global Positioning System (GPS) has changed the nature of human spatial navigation capabilities, with potential repercussions for the neural mechanisms that support spatial cognition (see Maguire et al. 2000).[5]
- **Mobile Computing and Wearable Devices:** Mobile access to the Internet is a key capability area for many technology vendors. A range of mobile Internet-enabled devices are currently available, of which the smartphone is undoubtedly the most popular. Mobility is a key factor in increasing our access to online data and services. It is also central to the emergence of what is sometimes referred to as the 'Embodied Web' (see Matsumoto et al. 2008)—the idea that our interactions with the Web will eventually occur as part of our everyday embodied engagements with a heterogeneous array of material artifacts. The fact that mobile devices, such as smartphones and wearable devices, are becoming so closely associated with the biological body, coupled with the fact that their modes of operation are increasingly linked to our physical movements and physiological processes, highlights the potential relevance of

---

[3] This represents a subset of all the areas that could have been listed. Other areas of notable interest from a cognitive science perspective include cloud computing (see Clowes 2015), the Semantic Web (see Smart, in press), and the Internet of Things (see Sect. Embedded Cognition).

[4] In particular, the use of linked data formats helps to separate issues of information presentation from issues of information representation. This kind of 'presentational agnosticism' is crucial when it comes to the flexible (and dynamic) creation of cues, prompts, and affordances that serve to shape the profile of human thought and action (see Smart, in press).

[5] It is also interesting to note the way in which this ubiquitous and ever-present 'data environment' helps to provide new opportunities for the implementation of location-aware intelligent systems, such as driverless cars and aerial drones.

the Internet (and Internet-enabled devices) to issues of material embodiment and embodied cognition (see Sect. Embodied Cognition).

- **Human Computation:** Human computation is an evolving area of research that seeks to harness human intelligence in the context of computational tasks, especially those whose complexity exceeds the capabilities of existing Artificial Intelligence (AI) algorithms (Michelucci 2013; Law and von Ahn 2011). Research in the area of human computation seeks to study the ways in which human capabilities can be integrated into some larger computational processing routine, yielding a form of bio-technologically hybrid computational system. In this case, the Internet often serves as a convenient platform for combining the complementary capabilities of conventional computing systems and human agents. Citizen science is one area where Web-based forms of human computation system are often used to good effect (Lintott and Reed 2013). Compelling examples of such systems can now be found in a number of scientific areas, including proteomics (Khatib et al. 2011), astronomy (Lintott et al. 2008), and neuroscience (Marx 2013; Helmstaedter et al. 2013). From the perspective of situated cognition, such research is relevant to issues of distributed cognition and the realization of bio-technologically hybrid forms of machine intelligence (Michelucci 2016; Smart 2013).
- **Augmented/Mixed Reality:** Augmented/mixed reality devices aim to 'embed' digital information in the real world by adding virtual overlays to the real-world environment. Some notable innovations in this area include Google Glass and Microsoft HoloLens. The cognitive impact of such devices has typically been discussed in relation to both embodied (Smart 2014) and extended cognition (Smart 2012); however, we can also view technological advances in this area as relevant to issues of embedded cognition (see Sect. Embedded Cognition).
- **Personal Informatics and Personalization:** Issues of personalization and the 'quantified self' (see Swan 2013) are increasingly popular focus areas for WAIS research. Although the majority of research in this area is concerned with personal data stores (Van Kleek and O'Hara 2014) and self-tracking technologies (Lupton 2013), there is also considerable interest in the cognitive and epistemic impact of techniques that support personalized access to online information (e.g., Simpson 2012). Issues of personalization and personal informatics are of interest from the perspective of both extended (Clowes 2015) and embodied cognition (see Sect. Embodied Cognition).

As is apparent from this (admittedly partial) list, our near-term future is one in which the Internet is likely to play an ever-more intimate role in shaping the nature of our everyday cognitive activities. Relative to this influence, it makes sense, we suggest, to see the Internet as a important part of the cognitively-potent extra-organismic environment in which our biological brains are now situated. The Internet, in other words, should be seen as a form of cognitive ecology. Not only does this perspective help to establish an important link with ecological approaches in contemporary cognitive science (Hutchins 2010; Malafouris 2013; Tribble and Sutton 2011), it also helps us to think about the creation and modification of online content as a form of 'ecological engineering' (Sterelny 2003) or 'cognitive niche construction' (Clark 2008). Every time we upload, annotate, edit, tweet or post, we are all (arguably) engaged in the construction and configuration of an environment that (in return) plays an increasingly important role in shaping the profile of our cognitive and intellectual endeavors.[6]

Situated Cognition Theory

In the last 20–30 years, there has been a shift in the cognitive sciences away from cognitive processes realized in the brain towards cognitive processes involving brain, body, and the environment (e.g., Clark 1997). In *The Cambridge Handbook of Situated Cognition*, Robbins and Aydede (2009) identify three distinct but related theses that characterize the situated cognition movement. First, the embodied cognition thesis, which claims that cognition depends on, and is sometimes constituted by, the human body (e.g., Gallagher 2005). Second, the embedded cognition thesis, which claims that our cognitive processes are sometimes shaped but not constituted by bio-external resources (e.g., Rupert 2004). Third, the extended and distributed cognition theses, which claim that cognitive states and processes, under certain conditions, are distributed across embodied agents and cognitive artifacts or other bio-external resources (e.g., Hutchins 1995; Clark and Chalmers 1998). Some theorists take these three approaches as a package deal, whereas others defend only one of these approaches. In this chapter, we do not prioritize one approach over another but see merit in all of these approaches as a means to better understand the cognitive implications of the Internet.

Situated cognition theory can thus be seen as the genus and embodied, embedded, and extended/distributed cognition theory as its species.[7] All these approaches have conceptual and metaphysical consequences, since they move beyond an individualist form of cognitivism and towards a picture that involves brain, body, and environment. As a result, these approaches also have methodological consequences: rather than merely focussing on cognitive processes in the brain, they advocate that we should focus on the cognitive relation between the

---

[6] Aside from its cognitive and epistemic effects, the Internet also influences the structure and organization of social processes. This raises a wealth of (socio-economic, socio-cultural and socio-political) issues that are the current focus of attention within the social sciences and digital humanities (Lupton 2015; van Dijk 2012). An in-depth discussion of such issues is beyond the scope of the current chapter; however, it is important to bear in mind that the Internet may sometimes be seen to exert an indirect influence on cognitive processes as a result of its ability to reshape the wider social, political, cultural and economic landscapes in which much of human thought and action takes place.

[7] Whilst this is a helpful tripartite distinction between the species of situated cognition theory, it is not exhaustive. There are other views such as enactivism (Stewart et al. 2010), collective cognition (Smart et al. 2010b) and transactive memory theory (Wegner 1995) that are also non-Cartesian in that they are concerned with the way a cognitive agent is situated in the environment.

brain, body, and environment, both on a conceptual and empirical level. Therefore, in order to study situated cognitive systems, Clark argues that we need “a new kind of cognitive scientific collaboration involving neuroscience, physiology, and cultural, social and technological studies in about equal measure” (Clark 2001, p. 154). One of the goals of this chapter is to contribute to Clark’s call by conceptualizing human-Internet interactions in light of the different species of situated cognition theory.

The relationship between these three species is not always clear, as there are different versions of each species with claims of various strengths. In the case of embodied cognition, for example, one can make a distinction between weak and strong embodied cognition. Weak embodied cognition claims that human cognitive processes sometimes depend on and are shaped by the body, but are not constituted by it. Strong embodied cognition, on the other hand, claims that cognition is partly constituted by the body. It seems that the embedded and extended cognition theses have to incorporate (some version of) the embodied cognition thesis. This is so because one needs to interact bodily with environmental structures and cognitive artifacts in order to scaffold or extend one’s cognitive processes. The body thus plays an essential role in embedding or extending cognition. However, one can be an embedded cognition theorist without accepting the strong version of the embodied cognition thesis. So, one can claim that cognitive processes depend on bodily processes and environmental structures, but deny that these are constitutive of cognition. Body and environment are then mere causal input to the brain-bound cognitive system. This relationship is different for extended cognition, which seems to include a strong version of embodied cognition as demonstrated by Clark’s phrases “putting brain, body and world back together again” (Clark 1997) and “where brain, body, and world collide” (Clark 1999). Here, the body is seen as one of the constitutive elements in a cognitive process that is distributed across brain, body, and the extra-organismic environment. Paul Dourish (2001), for example, is an embodied cognition theorist who is interested in how embodied creatures like ourselves interact with computational systems like tangible computing systems. He refers to this as ‘embodied interaction’. These kinds of computational systems, however, are not conceived of by Dourish as part of the human cognitive system. Instead, they are seen as resources that shape embodied cognitive processes. As such, Dourish is an embodied cognition theorist who embraces an embedded, but not extended, perspective.

Robbins and Aydede (2009) lump together the extended and distributed cognition theses, as both of these views argue that external resources can be constitutive of cognition. There are, however, important differences between the two views. Hutchins (2014) points out that extended cognition refers to a subset of cognitive events that involve the interaction of internal and external resources. Distributed cognition, by contrast, is a view on *all* of cognition. So the question is not whether or when cognition is distributed; rather “…the interesting questions concern the elements of the cognitive system, the relations among the elements, and how cognitive processes arise from interactions among those elements…The hypothesis of extended cognition is an important hypothesis *within the perspective* *of distributed cognition*” (Hutchins 2014, p. 36) [emphasis added]. Distributed cognition is thus a much broader view than extended cognition. Furthermore, extended cognition is ‘organism-centered’ in the sense that it sees the biological (human) agent as playing a crucial role in the assembly, maintenance and monitoring of extended cognitive circuits (see Clark 2008, p. 139). Distributed cognition theory, in contrast, does not assume that humans are necessarily the center of distributed systems. “Centers and boundaries are features determined by the relative density of information flow across a system”, suggests Hutchins (2014, p. 37). This marks an important difference between extended mind theorists and the proponents of so-called distributed cognition. Extended mind theorists, such as Clark (2008), emphasize the crucial role of the biological brain in the assembly of extended cognitive systems. This, as Clark notes, leads to the idea that cognition is organism-centered, even if it isn’t always organism-bound. Distributed cognition theorists, such as Hutchins (2011), tend to object to this organism-centered view. They worry that “by stressing the pivotal role of the brain in the recruitment of external resources and in the maintenance of resource-engaging cycles, [the extended mind theorist] actually gives too much away to a traditional internalist vision” (Clark 2011, p. 451).

Much of the research into distributed cognition has typically focused on socio-technical systems that involve the interaction of (multiple) human agents with non-biological props, aids and artifacts (Hutchins 1995). Given that many kinds of Internet-based systems are ones that feature a combination of social and technological elements, it seems that the Internet might be a natural place to look for instances of distributed cognizing. As already mentioned, however, the use of the term ‘distributed cognition’ is apt to cause confusion. It might seem, for example, that by using the term distributed cognition we are referring to a particular form of cognition, one that is, in this case, realized by multiple individuals acting in concert with an array of non-biological props, aids and artifacts. Recently, however, Hutchins (2014) has suggested that this is the wrong way to view the notion of distributed cognition. “Distributed cognition,” he suggests, “is not a kind of cognition; it is a perspective on all of cognition” (p. 3). The term ‘distributed cognition’ is thus equally applicable to cases of brain-based cognizing and to forms of cognition that occur in the context of larger hybrid ensembles consisting of both biological and non-biological elements. In view of this, we suggest that the term ‘collective cognition’ is a more appropriate way to characterize the cognitive processing that occurs in socially-distributed contexts, irrespective of whether or not the relevant information processing loops involve technological or artifactual mediation. Such a move may seem undesirable, especially if one regards the term ‘collective cognition’ as referring solely to systems in which multiple individuals participate in face-to-face exchanges. Giere (2007, 2012), for instance, attempts to make a distinction between what he calls “full-blown distributed cognition” and “mere collective cognition” (Giere 2007, p. 1). In particular, he suggests that systems in which instruments and other artifacts form parts of a cognitive system should not be seen as instances of collective cognition. Instead, he suggests that we should limit the term ‘collective cognition’ to situations in which multiple human agents are working together. It will be clear, however, that the way in which Giere (2007, 2012) interprets the term 'distributed cognition' is in some conflict with the view adopted by Hutchins (2014). In addition, we should perhaps question the extent to which it is actually possible to distinguish between cases of distributed and collective cognition in the manner suggested by Giere. One potential problem, here, is that even in face-to-face contexts it does not seem entirely appropriate to grant that *no* form of artifactual mediation is taking place. Human social exchanges are typically mediated by linguistic symbols of either the verbal or orthographic variety, and these sorts of symbolic representations are sometimes referred to as artifacts. Clark (1997, Chap. 10), for example, suggests that we should see public language as the 'ultimate artifact', as a kind of tool that "enables us to reshape a variety of difficult but important tasks into formats better suited to the basic computational capacities of the human brain" (p. 193). In view of all this, we will use the term 'collective cognition' to refer to systems of multiple agents that participate in cognitive tasks (i.e., tasks that are typically glossed as cognitive in nature), irrespective of whether or not these systems involve technological or artifactual mediation.

Having thus briefly analyzed some of the relations between the species of situated cognition theory, we now point out two important topics in situated cognition theory, namely (1) the size of the unit of analysis and (2) whether an external resource is constitutive of a cognitive state, process or system. Some theorists focus on small scale systems comprising a single embodied agent interacting with a cognitive artifact (Clark 2008; Kirsh and Maglio 1994), or two people interacting with each other, thereby forming a transactive memory system (Sutton et al. 2010; Wegner 1995). Others focus on larger systems consisting of many embodied agents interacting with a number of artifacts such as a team of navigators on a ship (Hutchins 1995), a group of scientists working in a laboratory (Nersessian 2009), or larger social groups such as sport teams (Williamson and Sutton 2014; Theiner et al. 2010). In this chapter, our units of analyses are both single embodied agents interacting with the Internet as well as collections of such agents. We also look at systems where online information is merely scaffolding cognition and where it is potentially integrated deeply into the cognitive processes of their users, thereby extending their cognitive processes.

Embedded Cognition

Unlike the notion of extended cognition, which sees elements of the extra-organismic environment as sometimes playing a constitutive role in the realization of cognitive states and processes, the notion of embedded cognition rejects the idea that the boundaries of cognition extend beyond the traditional biological borders of the human agent. Embedded cognition theorists thus focus on how artifacts and other external resources aid, but do not constitute, our cognitive systems (e.g., Rupert 2004). In this subsection, we look at how the Internet aids and shapes our cognition.

Throughout the evolution of our cognitive system, our minds have depended heavily on material culture (Gregory 1993; Vygotsky and Luria 1994; Donald 1993). Lambros Malafouris (2004, 2013) argues that tools and other human-made objects are the conditions of possibility for a range of cognitive operations which we often too carelessly attribute to our biological heritage. The invention of written language, calculation devices, maps, the printing press, and other cognitive technologies have augmented and shaped our cognitive capacities. The Internet, on this view, is just the latest in a long history of cognition shaping and enabling tools, opening some cognitive spaces and perhaps closing others.

Due to the Internet, we are currently living through important changes in the material (and digital) culture of memory, as an increasingly varied range of digital devices provide E-Memory adjuncts to, and extensions of, our biological resources. E-Memory can be defined as digital systems, devices and services that “we use to record, store and access digital memory traces to augment, re-use or replace organismic systems of memory” (Clowes 2013, p. 107). Research in cognitive psychology appears to show that human cognition is already factoring in the presence of E-Memory resources. In cases where a computer file is believed to remain available in the future, there is a tendency for our biological memory systems to remember the location rather than the contents of the file (see Sparrow et al. 2011). If this is a general indication of how our brains deal with ambient information stores, it is likely they are already being restructured by constantly accessible online systems. It is still undetermined how much our biological memories are affected by the cognitive affordances of the Internet and much more empirical data is needed. However, it is clear that we are becoming ever-more accustomed to using search technologies to recall facts and settle arguments. As we lean ever-more heavily on these technologies our basic habits of mind and thinking are undergoing significant changes.

When some theorists (e.g., Carr 2010) bemoan that human cognition is under threat from a host of new Internet technologies, they tend to forget that it is not a pristine human mind which is under threat from technology, but rather an artifact-dependent mind faced with the need to accommodate a new set of technologies (Donald 1993). However, it is clear that E-Memory has a range of properties which make it very different from previous epochs of memory technology. For this reason, E-Memory potentially has novel implications for our cognitive systems (Clowes 2013).

Using previous external memory systems such as tying knots in a string, making marks in clay, or maintaining written records tended to be effortful tasks and ones where storage capacity was finite (Mayer-Schönberger 2011; Donald 1993). E-Memory, in contrast, can record vastly more than previous regimes of technology. In principle, we can now store and retrieve a lifetime’s worth of high resolution video and sound (Gemmell and Bell 2009; Gemmell et al. 2006). Gordon Bell calls this phenomenon “total recall”. Additionally, recording, and to an extent accessing, E-Memory traces has become a relatively effortless task for an individual equipped with a smartphone. This suggests that Internet-mediated E-Memory will have a very different informational profile to past memory technologies. As we come to efficiently and unthinkingly rely upon the Internet, the human cognitive profile is likely to undergo significant changes. Much will depend on how our flexible cognitive architecture accommodates to the new environment. The uses of E-Memory we have so far discussed are largely mediated by personal mobile devices we carry with us. The growing trend, however, is for the wider material environment to become augmented with Internet-mediated technologies. We are deliberately re-designing the material world to be “smarter”.

Given that embedded cognition recognizes the role of extra-organismic factors in shaping human cognitive processes, it should come as no surprise that there are many points of interest for the embedded cognition theorist attempting to understand the cognitive implications of the Internet. In fact, practically all the features of the Internet ecology that surface in relation to embodied, extended and collective cognition also come into focus when we view the Internet through the lens of embedded cognition. There are, however, a number of specific points of interest for the embedded cognition theorist. These relate to the way in which the Internet of Things (IoT) and augmented reality technologies are able to alter the features of the physical (and virtual) environments in which we human agents are materially embedded.

The aim of the IoT initiative is to equip a variety of everyday physical objects with data acquisition, data processing and data exchange capabilities (Greengard 2015; Miller 2015). In addition, the IoT tends to view the environment as something of a cooperative partner with respect to the performance of a multitude of different tasks. Crucially, the IoT promises to deliver the kinds of capabilities that are often alluded to in discussions of ubiquitous computing (Weiser 1991) and ambient intelligence (Weber et al. 2005). The general idea is that by extending the reach of the Internet to the elements of our physical environment, we are able to transform the environment into something that is highly responsive to our needs and supportive of our thoughts and actions. This is the guiding principle behind current work that seeks to develop a seemingly endless array of ‘smart things’. These include, for example, smart TVs, smart cars, smart buildings, smart cities, and, of course, smart environments (see Miller 2015). From the perspective of embedded cognition, the IoT promises to alter the nature of our interactive engagements with the external environment, thereby influencing the kinds of dependencies that are deemed to shape brain-based forms of cognitive processing. In particular, we can see the advent of the IoT as part of an attempt to structure the environment in ways that enhance our biologically-based capabilities.

As a concrete example of the way in which the IoT may help to shape human cognition, consider the case of prospective memory. Prospective memory is a form of memory that involves “remembering to carry out intended actions without being instructed to do so” (Baddeley et al. 2009, p. 343). The case of an individual who needs to remember to defrost the meat by removing it from the freezer when they return home from work serves as a typical example of prospective memory. Such forms of memory are, of course, relatively commonplace, and they are pretty much indispensable in terms of our ability to coordinate our lives effectively—a fact that is all too sadly evidenced by those suffering from impairments in prospective memory (see Woods et al. 2008). As has been pointed out by a number of commentators (e.g., Staley 2014, pp. 36–37), the advent of smart environments provides a range of opportunities to reshape the nature of prospective memory. Staley (2014), for example, talks of smart devices being used to implement prospective memory systems that allow individuals to 'embed their intent' within specific environments, such as within their home or office. An individual could thus be reminded of the need to engage in particular actions (e.g., to remove the meat from the freezer) whenever they are suitably placed to perform these actions (e.g., when they first enter the kitchen upon returning home from work).[8] Prospective memory is also embedded in the mobile devices that we carry around with us. We can expect these devices to exhibit ever-greater levels of contextual sensitivity, e.g., reminding us to check our shopping list as we pass a grocery store. Relying on such devices to structure our activities can be seen as partially outsourcing human agency to our devices. However, this need not be seen as undermining our agentive powers. Instead, it can be seen as continuing a long history of using devices to structure, support and (even) generate complex aspects of human agency (see Neumann and Cowley 2013, for further discussion).

Aside from the IoT, another focus of interest for the embedded cognition theorist relates to the development of augmented and mixed reality devices. These devices (the exemplar of which is Google's Project Glass) support the creation of virtual overlays that are superimposed on the real-world environment. A crucial point of interest, here, concerns the way in which such devices can be seen to modify the properties of the environment in which cognition occurs. By generating an array of virtual representations, augmented reality devices are able to alter the 'effective' structure of the local environment, expanding the array of informational cues and affordances that can be used to guide cognitive processing. The inherent flexibility of these 'virtual designer environments'[9] means that future generations will have an unprecedented opportunity to rapidly reconfigure the structure of their environments in ways that complement, supplement or perhaps even supplant their brain-based cognitive capabilities.

---

[8] Note that inasmuch as we see prospective memory as a form of memory in which we perform future actions *without explicit instructions* (see Baddeley et al. 2009), it is unclear to what extent we should regard reminder systems as implementing a form of (external) prospective memory.

[9] This notion of a 'virtual designer environment' builds on the notion of a 'designer environment' as discussed by Clark (1997): "We build 'designer environments' in which human reason is able to far outstrip the computational ambit of the unaugmented biological brain" (Clark 1997, p. 191).

Embodied Cognition

Although there are a number of different views as to what is implied by the term 'embodied cognition', a common feature of embodied cognition research is the emphasis that is placed on extra-neural bodily factors in shaping the course of cognitive processing (Anderson 2003; Shapiro 2007, 2011). Typically, research into embodied cognition emphasizes the way in which an organism's bodily structure or physical actions help to constrain (and sometimes constitute) cognition. A somewhat trivial example is provided by the way in which the placement of an organism's sensory apparatus (the position of their eyes and hears) helps to structure the incoming sensory array in ways that support perceptual processing (Webb 1996). Other research focuses on the ways in which dynamically evolving motor state variables can help to guide the expression of intelligent behavioral responses (e.g., McBeath et al. 1995). More complex forms of embodied cognition research come in the form of work that seeks to evaluate the role of physical actions (e.g., hand gestures) in supporting various forms of human cognitive competence (Goldin-Meadow 2003).

At first sight, it might appear that a discussion of embodied cognition is somewhat out of place in a chapter that focuses on the role of the Internet in shaping our cognitive profile. After all, work in embodied cognition tends to focus on situations in which we are actively engaged with the real world, exploiting all manner of sensorimotor cues in order to realize intelligent thought and action. The nature of our interaction with the Internet seems far removed from this sort of situation. Although we might be justified in seeing the Internet as an important part of the context in which cognition occurs—part of the material backdrop against which our thoughts and actions take shape—it is by no means clear that the details of our physical embodiment really matter that much when it comes to understanding the cognitive consequences of our online interactions.

There are, in fact, a number of ways in which the Internet implicates issues that lie at the heart of the embodied cognition research programme. Firstly, as noted by Smart (2014), the advent of mobile and portable computing solutions is progressively altering our sense of what it means to engage with the online world. In place of conventional forms of interaction, in which we interact with the Internet via a browser interface while seated at a desktop computer, it is increasingly common for us to engage with the Internet as part of our embodied interactions with the wider physical environment. Mobile devices, such as smartphones, for example, enable us to interleave our interactions with the Internet and the real world in a way that seems to blur the traditional distinction between 'offline' and 'online' modes of interaction (see Floridi 2011, 2014). In addition, as new kinds of Internet-enabled device become available, so the palette of physical actions and gestures that we use to interact with the Internet is expanding. Touchscreens have clearly played an important role, here, with swiping and zooming emerging as more-or-less standard parts of our gestural lexicon. Other kinds of interactivity aim to capitalize on the way in which we typically interact with a common array of physical artifacts and objects, helping to support forms of 'embodied interaction' (see Dourish 2001) with the online world. Consider, for example, work by Matsumoto et al. (2008) to develop a Web-enabled umbrella. The umbrella features a variety of sensors (e.g., GPS, compass, accelerometer, etc.), and it is able to project Web-based content directly into the user's field of view by virtue of a projection device focused on the underside of the umbrella canopy. By providing the user with a range of interaction opportunities (e.g., the normal turning, dipping, and twisting actions that people perform with umbrellas) and by also integrating information from a variety of sensors and Web services, the umbrella is able to present a variety of forms of context-relevant information that take into account both the user's physical location, as well as their current interests and activities. Of particular interest in the current context, Matsumoto et al. (2008) describe their work as part of an effort to realize what they call the 'Embodied Web': a form of enhanced interactivity in which natural embodied interactions are used to interact with the Web and "make our experience in the real world more engaging and active" (Matsumoto et al. 2008, p. 49).

A second way in which notions of embodied cognition are relevant to understanding the cognitive significance of the Internet comes in the form of what might be referred to as 'the extended body'. The idea, here, is that (in certain cases) it may make perfect sense to see Internet-enabled devices as literal prosthetic extensions of an individual's biological body. If this idea seems to strange or unpalatable, perhaps it will help to reflect on what it is that makes something a genuine body part. One possible answer to this question focuses on the way in which our bodies work to mediate our sensorimotor engagements with the world. Our ears therefore count as part of our body because they assist with the transduction of certain kinds of energetic fluctuation in the ambient environment; our legs count as part of our body because of the way in which they service our locomotory objectives; and our teeth count as part of our body because of the way in which they enable us to physically prepare certain kinds of matter for the processes of digestion and absorption. Note that an appeal to the biological nature of the candidate bodily elements will not suffice here. Should a cochlear implant, a prosthetic limb or a tooth implant fail to count as part of our body simply because they are not biological in nature? And what about the 'bodies' of smart cars and mobile robots?[10] Do these systems fail to have a body simply because they are not biological systems? The answer to these questions is surely a resounding 'no'. And once we drop appeals to biology in determining what it is that makes something a part of the body then the path is clear for a more functionally-oriented conception of the body. Crucially, once we have this functional view to hand, it becomes possible to consider non-biological resources as literal extensions of the biological body: providing a non-biological

---

[10] Importantly, issues of embodiment often surface in the context of research into cognitive robotics (see Pfeifer and Bongard 2007). This highlights the importance of a non-biological conception of the body to embodied cognitive science: in the absence of such a conception it becomes difficult to adopt a unified perspective of research into a rich variety of materially-diverse (e.g., biological, robotic and virtual) embodied cognitive systems (see Smart and Sycara 2015).

element plays the same sort of functional role as a conventional (biological) body part, then it seems we should treat that non-biological element as a genuine part of the body. This, of course, opens the door to cognitively-potent forms of 'corporeal incorporation' involving a variety of non-biological resources.

One example of a functionally-oriented conception of the body is provided by Clark (2007, 2008). He suggests that we should identify the body with whatever it is that just so happens to serve as the "locus of willed action, the point of sensorimotor confluence, and the stable (though not permanently fixed) platform whose features and relations can be relied upon (without being represented) in the computations underlying some intelligent performances" (Clark 2008, p. 207). The claim, in essence, is that we should identify the body with whatever it is that is playing the sort of role that our biological body typically plays with respect to the genesis and organization of intelligent behavior. Inasmuch as we accept this claim, then it seems that forms of bodily extension that involve our current arsenal of portable and mobile Internet-enabled devices are a realistic possibility. In other words, there seems to be no principled reason why Internet-enabled devices should not be counted as, on occasion, functioning as literal body parts. The only question, of course, is to what extent such devices actually are apt for bodily incorporation. This is an issue that must appeal to the nature of our interaction with Internet-enabled devices, specifically the extent to which such forms of interaction satisfy the sort of functional criteria alluded to by (e.g.) Clark (2008). If we accept that Internet-enabled devices do actually function as literal body parts—as prosthetic technological extensions that enable us to sense, manipulate, exploit, and alter the *online* world—then there seems to be no clear reason why such devices should not be of relevance and interest to embodied cognitive science.[11] This is because embodied cognition is concerned with the way in which bodily forces and factors influence cognition. When technological resources become bodily protheses they may also (*qua* embodied cognition) function as cognitive prostheses. This is particularly so when such prostheses are recruited as part of our cognitive and epistemic endeavours.

A variety of sources of (largely circumstantial and anecdotal) evidence support the idea that issues of bodily extension may be relevant to a range of Internet-enabled devices, most notably those that we carry around with us (e.g., smartphones), or those that we attach to the biological body (e.g., smartwatches). Recent trends in technology development are thus largely consistent with what Biocca (1999) refers to as 'progressive embodiment', the idea that technological advances entail the "steadily advancing immersion of sensorimotor channels to computer interfaces through a tighter and more pervasive coupling of the body to interface sensors and displays" (Biocca 1999, p. 5). Irrespective of whether we

---

[11] The thing that is important to remember, here, is that inasmuch as a non-biological resource counts as part of an organism's body, then (relative to the claims made by proponents of embodied cognition) the resource is (potentially) poised to play a role in shaping that organism's cognitive processing routines. As a result, if an Internet-enabled device counts as a part of the body (on the basis of functional criteria), then it seems that it should be just as much a focus of analytic attention for the proponent of embodied cognition as should a more conventional (i.e., biological) body part.

accept that the boundaries of the human mind are moving outwards (see Clark and Chalmers 1998), it appears that our technologies are progressively reaching inwards, attempting to establish ever-more intimate associations with the elements of the biological body (see Lynch 2014). Our emerging panoply of portable devices are thus not just forms of smart machinery, they are also (potentially at least) forms of 'intimate machinery'.

The mobile phone has, of course, been a popular focus of attention when it comes to issues of bodily extension. Drain and Strong (2015), for example, suggest that the smartphone "becomes incorporated within the assemblage of bodily appendages, environmental features, and artifacts that we encounter in everyday life, to the point where the phone can be considered as a prosthetic extension of ourselves" (p. 190). A number of studies have also revealed that users often regard their mobile phones as extensions of their 'self' or body (Oksman and Rautianen 2003a, b; Gant and Kiesler 2001). Perhaps such results should not come as a surprise given the way in which many individuals now relate to their mobile phones. To an ever-greater extent, the smartphone is an indispensable instrument that enables the individual to negotiate the various social, cognitive and epistemic challenges that they confront as part of their daily life (Drain and Strong 2015; Miller 2014). This is often reflected in the kinds of deep emotional attachment that people have with their mobile devices (Miller 2014). As Vincent et al. (2005) note "for some people [their mobile device] has become almost an extension of their body as they hold and fondle the device even when the device is not in use" (p. 72).

One further line of research that may be relevant to claims regarding the bodily incorporation of Internet-enabled devices comes from a study by Salerno et al. (2012). Salerno et al. (2012) sought to investigate the neural processing of self- and other-related stimuli using a trans-cranial magnetic stimulation technique. As part of their experimental protocol, human subjects were presented with four kinds of images. These showed (1) the subject's own hand, (2) the hand of another human subject, (3) the subject's own mobile phone, or (4) the mobile phone of another person. Interestingly, Salerno et al. (2012) observed similar neurological responses when subjects were presented with images of both their own hands *and* their own mobile phones (i.e., self-related stimuli). However, these responses were distinct from those elicited by the images of 'other-related' stimuli (i.e., the hands or phones of other people). Although the psychological significance of these results is unclear at the present time, Salerno et al. (2012) note that issues of bodily extension may be relevant to their findings. Interestingly, the effects observed by Salerno et al. (2012) were specific to the right hemisphere of the brain. This is consistent with neuropsychological research indicating that the right brain hemisphere is important to a sense of body ownership, with damage to the right hemisphere resulting in a form of 'disownership' of particular body parts (Aglioti et al. 1996; Vallar and Ronchi 2009). Aglioti et al. (1996), for example, report the case of a woman with damage to the right hemisphere who denied ownership of her left hand and associated 'extracorporeal' objects, for example, the rings that were typically worn on her left hand. Summarizing their results, Aglioti et al. (1996) suggest that "the mental image of one's body may include inanimate objects which [have] been in contact or in close proximity with the body itself" (p. 293). They go onto claim that their findings can be interpreted as providing support for the notion of an 'extended body schema', in which extra-corporeal objects become incorporated into the neural processes that mediate a sense of one's body and its capacities for action.

In addition to issues of embodied interaction and bodily extension there is a third way in which notions of embodied cognition may prove useful in enabling us to track the cognitive impact of Internet-related technologies. The main point of interest, here, relates to the way in which new technologies are reshaping the kind of opportunities that we have to track our bodily states and actions. Wearable technologies thus provide not just a means to transform the kind of contact we have with the online world, they also transform the kind of relationship we have with ourselves (see Lupton 2015). Crucially, as wearable technologies become more sophisticated, they often become more capable of sensing physiological and behavioral states that enable them to contextualize their modes of operation in novel ways. Many contemporary devices already feature this kind of context sensitivity. For example, when I am located in New York, and I wish to visit a particular museum, I can rely on the GPS capabilities of my Internet-enabled device—an iPhone let's say—to adapt its information retrieval processes in a manner that is germane to my current interests. The device, we can assume, is sensitive to my geographic location in a way that enables it to return information that is likely to be directly relevant to my immediate needs and concerns. A similar form of context sensitivity is likely to become possible with future wearable technologies. In this case, however, the devices are likely to factor in a far richer range of cues and affordances than is possible with today's technology. Importantly, by being able to monitor body-related information it becomes possible for this new suite of devices to sensitize their operation to factors that are more directly relevant to our cognitive and epistemic activities. Consider, for example, work by Koriat and Nussinson (2009) to investigate the physiological correlates of the feeling of knowing. They report that the tension of the corrugator muscle can be used to detect the subjective experience we have when we feel we know something. Inasmuch as future devices are able to track physiological signals that index a variety of epistemic feelings (such as the feeling of knowing or the feeling of difficulty) (see Michaelian and Arango-Muñoz 2014), it is possible that future technologies may be able to adaptively modify their modes of operation to support human end-users with regard to a variety of epistemically-related activities.[12]

The use of technology to record or track personal information is a central element of work that goes under the heading of the 'quantified self' (Lupton 2013;

---

[12] A similar point is made by Kunze et al. (2013). They suggest that the use of mobile sensing technologies portends an era in which technology is able to recognize and monitor various forms of cognitive activity, revolutionizing our understanding of the factors that contribute to optimal cognitive performance, as well as providing new ways for technology to shape and scaffold our cognitive routines.

Swan 2013). This is a term that is used to refer to any form of self-tracking activity, where the information that is typically tracked is of a biological, behavioral or environmental nature. Current forms of self-tracking include the recording of body weight, energy levels, time usage, sleep quality, health, cognitive performance, athletic performance, and learning strategies (Swan 2013). Clearly, not all these forms of self-tracking are ones that need to rely exclusively on Internet-enabled devices; nevertheless, an increasing number of devices are becoming available that attempt to automate or facilitate self-tracking efforts. As Swan (2013) notes:

> An increasing number of new personal data streams are being generated through quantified self tracking devices, biosensors, wireless Internet of Things devices, health social network data, and social media data. Additional personalized data streams from consumer EEGs, eye-tracking and emotion measurement could be coming in the future (p. 91).

The significance of Internet connectivity, here, is that it enables personal information to be stored online and subjected to forms of online (e.g., cloud-based) manipulation and processing. Such forms of online processing are often essential when it comes to the analysis of personal data, especially when one considers the quantity of data that can be generated by (e.g.) heart rate sensors. It should also be relatively clear that by placing information online, personal data is often made more amenable to large-scale forms of statistical analysis that can provide potential insights into (e.g.) health-related phenomena. The public availability of self-tracking data is thus a potential boon to epidemiological and clinical research, supporting analyses into factors related to disease onset and progression.[13]

One of the implications of the quantified self is that it provides a greater degree of awareness regarding one's bodily states and processes. Self-tracking technologies are thus sometimes seen as a means of creating a digital dashboard for the biological body, enabling individuals to tap into a wealth of previously inaccessible data. Some writers thus talk about self-tracking devices as supporting the emergence of technological 'exosenses' that extend the reach of the body's sensory capabilities:

> …the quantified self may become additionally transformed into the extended exoself [i.e., a suite of exosenses] as data quantification and self-tracking enable the development of new sense capabilities that are not possible with ordinary senses. The individual body becomes a more knowable, calculable, and administrable object through QS [quantified self] activity, and individuals have an increasingly intimate relationship with data as it mediates the experience of reality (Swan 2013, p. 85).

The general idea, therefore, is that self-tracking affords a new way in which body-related information can come to influence the course of cognitive processing. In particular, by virtue of their ability to make body-related information explicitly accessible and perceptible through the other senses,[14] the vision of the quantified

---

[13] For example, personal data could be used to evaluate hypotheses concerning the links between sleep patterns and the etiologic bases of diabetes (Tasali et al. 2008), depression (Landsness et al. 2011) and dementia (Sharma et al. 2015).

[14] This is sometimes referred to as para-synthetic expression (Won et al. 2015).

self opens the door to forms of embodied cognition in which issues of technological mediation are of critical importance. It has long been known, for example, that individual cognitive and emotional responses can be shaped by the explicit perception of physiological signals (e.g., Valins 1966). In view of this, it seems entirely possible that future technologies could play an important role in determining the role of bodily states and signals in influencing our everyday thoughts and actions at both an individual and social level (see Janssen et al. 2010).[15] This is, of course, an issue that lies at the heart of contemporary work in embodied cognitive scienceembodied cognitive science.

Extended Cognition

Having looked at how the Internet aids some of our cognitive processes and how it relates to human embodiment, we now continue by conceptualizing how it may extend our cognitive states and processes. Extended mind theory takes a systems perspective on the relation between a cognitive agent and the social and material environment in which that agent is situated. In essence, the extended mind is a claim about the location of the material substrate on which cognitive states and processes supervene. It is thus a metaphysical claim with epistemological and methodological consequences.

John Sutton (2010) distinguished between two waves of extended mind theorizing. The first wave argues for the parity of internal and external resources (Clark and Chalmers 1998), whereas the second wave argues for the complementarity of internal and external resources (Heersmink 2015, 2016; Sutton 2010). Sutton argues that external resources are often not on a par with internal resources. In the case of memory, for example, internal memory systems are dynamic, integrative and subject to degradation of their traces, whereas external information is often static, not integrated with other information and not subject to degradation. The complementarity view, as formulated by Sutton, reads as follows:

> In extended cognitive systems, external states and processes need not mimic or replicate the formats, dynamics or functions of inner states and processes. Rather, different components of the overall (enduring or temporary) system can play quite different roles and have different properties while coupling in collective and complementary contributions to flexible thinking and acting (Sutton 2010, p. 194).

Relative to this view, we can see that external resources need not replicate internal resources; instead, they can have different properties, functions and formats to those of the biologically-based cognitive system. It makes much more sense to rely upon and recruit external resources when they do something *different* from our native or biological capacities. Jointly, an embodied brain plus a complementary

---

[15] Explicit access to physiological information (e.g., heart rate) can also, on occasion, influence our sense of body ownership concerning a non-biological appendage (see Suzuki et al. 2013).

resource is a much more versatile and powerful cognitive system than just an embodied brain without any external aids. Wilson and Clark (2009) also defend a complementarity view. They write:

> Contrary to any requirement of fine-grained similarity then, what the friends of extended cognition actually expect, and study, are hybrid processes in which the inner and the outer contributions are typically highly distinct in nature, yet deeply integrated and complementary (Wilson and Clark 2009, p. 72).

The cognitive integration of an embodied agent and a resource with complementary properties and functions is thus important for better understanding extended cognitive systems. However, most of the current discussion on extended cognition and the Internet focuses on whether the Internet satisfies Clark and Chalmers' (1998) conditions of reliability, trust, accessibility and past endorsement (Clark 2008; Smart 2012, in press; Ludwig 2015). These conditions are often referred to as the "trust and glue" criteria, and they are typically seen as part of the first wave of extended mind theorizing. The Internet is often seen as failing to meet these criteria.[16] Clark and Chalmers (1998), for example, suggest that "The Internet is likely to fail on multiple counts, unless I am unusually computer-reliant, facile with the technology, and trusting, but information in certain files on my computer may qualify" (p. 18). In a later, co-authored publication, however, Clark says: "Perhaps external representations on the Web, when integrated appropriately into the processes that govern an agent's behaviour, may count as part of that agent's cognitive architecture" (Halpin et al. 2010, p. 2).

So, to better understand human-Internet interactions in terms of extended cognition theory, we need to focus on what "integrated appropriately into the processes that govern an agent's behaviour" means. We suggest that cognitive integration between an agent and the Internet is a multidimensional phenomenon. Some of the relevant dimensions include the kind and intensity of information flow between an agent and the Internet, the accessibility of information, the durability of agent-Internet couplings, the amount of trust a user puts into online information, the degree of transparency-in-use that is encountered, the ease with which the information can be interpreted, the amount of personalization that has been undertaken, and the amount of cognitive transformation introduced as a result of the

---

[16] Problems with trust often lie at the root of these concerns. Clark (2010), for example, claims people do not trust online content to the same extent that they trust information retrieved from bio-memory. From an empirical perspective, however, it is far from clear that people really do subject online information to the sort of evaluative scrutiny that would undermine its candidacy for cognitive incorporation (see Smart, in press). In addition, there a variety of reasons to suspect that at least some sources of online content can be implicitly trusted. Individuals may, for example, rely on the use of cloud-based personal data stores (see Van Kleek and O'Hara 2014) as a source of trusted information. They may also exploit a range of so-called 'online reliability indicators' (Smart and Shadbolt, in press) to guide metacognitive processes relating to information selection and endorsement (Arango-Muñoz 2013). Interestingly, processes that give rise to these indicators can, on occasion, be cast as a form of collective or distributed cognition. Ben-Naim et al. (2013), for example, present a distributed approach to the construction of (social) trust metrics, which are subsequently used to guide decisions relating to the endorsement of expert recommendations.

bio-technological merger (Sterelny 2010; Heersmink 2015, 2016; Menary 2010; Sutton 2006). The way we interact with online information may vary along all of these dimensions; however, the higher an agent-Internet system ranks on these dimensions, the denser the integration between the agent and the online information, and the easier it becomes (we suggest) to regard the agent-Internet system as engaging in a form of extended cognition. In this respect, it is interesting to note that the general thrust of technology design seems to be largely in favor of the emergence of cognitively-potent forms of bio-technological merger (see Smart, in press). Of particular interest, are a range of technologies that target our mnemonic capabilities (Clowes 2015). We thus use our mobile devices to store a lot of personalized information in the cloud and on the Internet, including appointments, birthdays, shopping-lists, sketches, annotated documents, to-do lists, notes, reminders, bookmarked timetables, and so on. The kinds of technologies that are used to interact with such online sources of information (e.g., mobile, portable and wearable devices) arguably serve to enhance the intensity of information flow between agent and the Internet, the accessibility of information, the amount of trust a user puts into the information, the degree of transparency-in-use, the ease with which the information can be interpreted, and the amount of personalization. For these reasons, the information is much more deeply integrated into the cognitive processes that govern our behavior, and it is therefore easier to see it as part of an extended cognitive system.

The way in which information is represented on the Internet may be of particular relevance when it comes to understanding the profile of human-Internet interactions. In particular, the transition to what has been referred to as the Web of Data (see Sect. The Internet: A New Kind of Cognitive Ecology) may play an important role in enabling Internet resources to be more closely integrated into everyday cognitive processing routines. Smart (in press, 2012), for example, suggests that the transition from a Web of Documents to a Web of Data plays an important role in enabling the Web to function as a component of bio-technologically hybrid cognitive systems. One reason for this concerns the accessibility of specific items of information—the fact that it is possible to retrieve isolated pieces of information in a wide variety of different task contexts. Another reason relates to the fact that content on the Web of Data becomes much more amenable to machine-based processes that can find, filter and format data in ways that are optimally suited to a human end-user's specific information needs and concerns (this is deemed to enhance the accessibility and functional poise of online information). Finally, the move away from the Web of Documents to the Web of Data opens up a range of presentational capabilities that can be used to guide thought and action in particular ways. Thus, when we think of the cognitive impact of new devices, such as Google's Project Glass and its successors, we should not necessarily think of their presentational capabilities as being limited to the display of conventional Web pages. Instead, we should think of a whole variety of different data-driven presentational capabilities, some assuming the form of simple natural language statements and instructions, others relying on the use of graphical cues, prompts and affordances. In addition, the notion of augmented, mixed or blended reality enables us to think of Internet-based information being used to create virtual overlays on the physical environment, enriching the kinds of environmental structures to which our brain-based processing routines are already attuned (see Sect. Embedded Cognition).

Collective Cognition

A recent focus of interest for the WAIS community relates to the use of Web technologies to support socially-distributed forms of cognition (Chi 2008, 2009; Kearns 2012); i.e., forms of cognition in which the relevant cognitive processes (e.g., reasoning, remembering and problem-solving) are distributed across a collection of individuals. This interest is reflected in a wealth of research relating to online socio-technical systems. The conceptual landscape of the Social Web is thus littered with terms like social computing (Parameswaran and Whinston 2007), human computation (Law and von Ahn 2011; Michelucci 2013), collective intelligence (Bonabeau 2009; Malone et al. 2010; Halpin 2013), social machines (Hendler and Berners-Lee 2010; Smart et al. 2014), technology-mediated social participation (Kraut et al. 2010), and the global brain (Heylighen 2013). The purpose of these locutions is often to emphasize the way in which Web technologies can be used to harness the socio-cognitive potential of large numbers of physically-distributed individuals.

Because of the kinds of opportunities it affords for large-scale collaboration, information sharing, and the coordination of collective efforts, the Internet seems to be ideally suited to supporting cases of collective cognition. Perhaps the best example of such support is provided by tasks that involve some form of collaborative problem-solving or collective decision making (Chi 2009; Chi et al. 2008). In this case, the Internet is often used to support a form of ‘virtual team-working’,[17] in which a collection of geographically-distributed individuals rely on the network-mediated exchange of information as a means of coordinating their individual problem-solving efforts. Given the popularity of such forms of virtual team-working in a range of organizational settings (see Kanawattanachai and Yoo 2002), it is perhaps not surprising that this particular form of collective cognition has become a prominent focus of attention for those working in the disciplines of computer, cognitive and network science. Recent research has thus sought to understand the factors that regulate the performance of collective cognitive systems under a range of experimental conditions, with issues of network structure (i.e., the time-variant organization of communication network topology) emerging as a particularly important focus of empirical investigation (Kearns 2012; Mason 2013;

---

[17] Virtual team-working is a form of team-working that relies on the use of information and communications technology to support task-relevant forms of information exchange, information processing and inter-agent coordination (see Powell et al. 2004).

Mason and Watts 2012; Mason et al. 2008; Smart et al. 2010a). The kind of processing that is undertaken in the context of such systems is sometimes glossed as a form of 'socio-computational processing', a term that helps to highlight the important role that technologically-advanced computational systems play with respect to the mediation of social exchanges and the active processing of task-relevant information.[18]

The notion of virtual teams[19] thus serves as an important example of how our existing notions of collective cognition—for example, group (Theiner 2014) and team cognition (Cooke et al. 2007; Salas et al. 2011)—can be applied to the online world. Virtual teams do not, however, exhaust the reach of the concept of collective cognition when it comes to understanding various forms of Internet-mediated activity. In some cases, the processes by which specific informational ecologies come into being on the Internet are themselves also described in distributed or collective cognitive terms. Consider, for example, the way in which the linking behavior of Internet users yields a body of information that can be used to support the operation of Internet search engines. The most popular example in this case, is, of course, the PageRank algorithm, as used by Google Search. Here, the editing actions of countless numbers of Internet users serves as the analytic substrate for machine-based processes that seek to enhance the accessibility of online information in a way that is aligned with the interests and concerns of the human user community. Heintz (2006) suggests that we should see such activities as a form of distributed (i.e., collective) cognition, one that (presumably) supervenes on the actions of a very large number of human individuals. Such claims are likely to prove controversial; however, Heintz's (2006) analysis does help to highlight the way in which a combination of human action and machine-based processing can play an important role in creating and configuring the online environment. A similar point is made by Smart and Shadbolt (in press) in regard to the social construction of 'reliability indicators'. Smart and Shadbolt suggest that the individual actions of many thousands of individuals helps to shape the kinds of cues and affordances that are available to guide the epistemic evaluation of online content. What is crucial here is the sense in which the social and technological components of some larger systemic organization are working to help shape and structure the nature of the online cognitive ecology—an ecology that then plays a significant role in sculpting the profile of our individual and collective cognitive endeavours.

---

[18] Although it is easy to see such forms of processing as a relatively recent phenomenon, it is important to remember that technologically low-grade variants of socio-computational processing date back to at least the 18th century (see Grier 2013).

[19] A virtual team, in this case, is simply a collection of individuals that engages in a form of virtual team-working (see Powell et al. 2004). Crucially, nothing in this definition rules out the possibility that a virtual team could (at different points in time) also function as a real-world (or face-to-face) team. The result is that any form of (conventional) team cognition is also (potentially at least) a form of virtual team cognition. Consider, for example, how collaborative sensemaking technologies (e.g., Shrager et al. 2010; Toniolo et al. 2014) might be used to support the kinds of analyses undertaken by (e.g.) criminal investigators (Baber 2013).

Wikipedia is another example where one encounters this particular form of cognitively-potent 'ecological engineering'. In this case, the technical components help to shape and scaffold user contributions in such a way as to yield a rich and reliable source of information that is relevant to a number of epistemic activities (Fallis 2011). Wikipedia is, in fact, emblematic of a broad range of systems, sometimes glossed as 'social machines' (Smart et al. 2014; Hendler and Berners-Lee 2010), in which much of the online content is supplied or generated by human end-users. Even if we demur from the conclusion that the processes associated with such systems should be regarded as genuine cases of collective cognition, we can surely accept that such systems support a form of ecological engineering that is relevant to our individual and collective cognitive capabilities.

Inasmuch as systems like Google Search or Wikipedia count as cognitive systems, they are clearly systems that exist at a much larger scale than the kind of kind of cognitive systems that are the typical focus of distributed cognition research. This, by itself, however, need not be a cause for concern. As Hutchins (2014) himself suggests cognitive systems can exist at a variety of spatial and temporal scales. In the ecological context of the Internet, we can identify a variety of socio-technical systems that could be considered as potential candidates for cognitive scientific analysis. Such systems can be ranked according to their relative social size, or the number of individual human agents that typically participate in them. At the lower end of this scale, we encounter systems that can perhaps be broadly construed as virtual teams (see above). The size of these systems typically ranges from tens to hundreds of individuals. At larger scales, we encounter human computation systems[20] (see Sect. The Internet: A New Kind of Cognitive Ecology) like Foldit (Khatib et al. 2011) and Galaxy Zoo (Lintott et al. 2008).[21] These systems—and others like them (see Lintott and Reed 2013)—rely on the efforts of thousands of individuals in order to support the process of scientific discovery. Systems such as Wikipedia and Google Search exist at even larger scales. In the case of Wikipedia, for example, tens of thousands of individuals participate in the editing of online articles, and Google Search requires even larger numbers of individuals (hundreds of thousands to millions). Large-scale forms of social participation are often critical to the success of these systems: in the absence of large-scale social participation, for example, Wikipedia could not have the coverage it does, nor could it update its articles in a timely fashion. Both of these features are of obvious relevance to Wikipedia's status as a source of epistemically-significant information.

---

[20] A human computation system, recall, is a system that combines human and machine capabilities to perform complex computational tasks (Law and von Ahn 2011; Michelucci 2013; Quinn and Bederson 2011).

[21] In the case of Foldit, human pattern matching and spatial reasoning abilities are used to help solve the problem of predicting the three-dimensional structure of selected proteins (Khatib et al. 2011). Galaxy Zoo, in contrast, relies on human visual pattern recognition to detect and classify galaxies from large-scale astronomical image databases (Lintott et al. 2008).

Conclusion

It is tempting to view the Internet as a globally-interconnected repository of data and information. This view of the Internet is not necessarily incorrect; however, it does not really do justice to the many ways in which the Internet shapes our actions and capabilities at both an individual and collective level. The Internet is thus not just an alternative way of disseminating information in a manner akin to that accomplished by our traditional print media (see Carr 2010). Neither is the Internet merely "the next great extension of the 'external symbolic storage system' humans have developed since the beginning of civilization" (Staley 2014, p. ix). The Internet is, in fact, many things. It is an instrument of social change, a mechanism for coordinating disaster relief efforts, a platform for scientific discovery, and a potential breeding ground for the next generation of AI systems. It is also (perhaps less positively) a tool for Distributed Denial of Service (DDoS) attacks on corporate Websites, a marketplace for the exchange of illicit goods, a vector for the transmission of extremist ideologies, and the future field of battle for warring nation states. The Internet is, in essence, an *environment* in which all manner of processes can occur and upon which many kinds of capabilities are founded.

When our attention is directed to the realm of human cognitive activity, we suggest the Internet plays an important role in shaping the nature of our cognitive processes. The Internet, we claim, forms an important part of the extra-organismic environment that shapes, scaffolds, supports, sustains and perhaps even realizes our cognitive processing routines. It is, in essence, an important part of the wider cognitive ecology in which our biological brains are situated.

The value of a situated approach to cognition is that it helps us to appreciate the many ways in which the Internet can impact our individual and collective cognitive capabilities. Understanding the nature of this influence is crucial. Even the most ardent advocate of neurocentrism would no doubt be willing to accept that the kind of environment in which a biological brain is situated says a great deal about the kinds of capabilities it can realize. A situated approach to cognition can help to reveal the many ways in which the Internet may influence our cognitive profiles, either for good or ill. It also helps to focus our attention on the capabilities of cognitive organizations whose systemic boundaries are not necessarily those of the (biologically-based) individual agent (e.g., extended and collective cognitive systems). Finally, a situated approach helps to reveal the complex web of reciprocal influences that exist between the Internet and those who use it. Importantly, the cognitive ecology of the Internet is an environment of our own making. By virtue of our interactions with the Internet, we help to shape what the Internet is, what it can do, and what it may yet become. In this sense, we are all, as Sterelny (2003) suggests, 'ecological engineers': we are all engaged in a process of cognitive niche construction (see Clark 2008), actively involved in the construction and configuration of a bio-external nexus of material resources that helps to influence the course of our cognitive processing and define the limits of our cognitive capabilities.

Acknowledgements Paul Smart's contribution to this work was supported by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished as part of the International Technology Alliance (ITA) Project under Agreement Number W911NF-06-3-0001. Some of the ideas contained in the chapter were presented by Richard Heersmink at a seminar at the Center for Human Interactivity of the University of Southern Denmark. Richard would like to thank Sune Steffensen for the invitation and the audience for helpful feedback. Richard would also like to thank John Sutton and Alexander Gillett for helpful discussion on the Internet and cognition. Robert Clowes wishes to acknowledge the support of Portuguese Fundação para a Ciência e Tecnologia grant/BPD/70440/2010 that made possible his contribution to this paper. All authors would like to thank two anonymous reviewers who provided detailed feedback and suggestions on an earlier version of the chapter.

References

Aglioti, S., Smania, N., Manfredi, M., & Berlucchi, G. (1996). Disownership of left hand and objects related to it in a patient with right brain damage. *NeuroReport, 8*(1), 293–296.

Anderson, M. (2003). Embodied cognition: A field guide. *Artificial Intelligence, 149*(1), 91–130.

Arango-Muñoz, S. (2013). Scaffolded memory and metacognitive feelings. *Review of Philosophy and Psychology, 4*(1), 135–152.

Baber, C. (2013). Distributed cognition at the crime scene. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice*. London, UK: Springer.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Hove, East Sussex, UK: Psychology Press.

Bateson, G. (1972). *Steps to an ecology of mind*. London, UK: Jason Aronson Inc.

Ben-Naim, J., Bonnefon, J. F., Herzig, A., Leblois, S., & Lorini, E. (2013). Computer-mediated trust in self-interested expert recommendations. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice*. London, UK: Springer.

Biocca, F. (1999). The cyborg's dilemma: Progressive embodiment in virtual environments. In J. P. Marsh, B. Gorayska, & J. L. Mey (Eds.), *Humane interfaces: Questions of method and practice in cognitive technology*. Amsterdam, The Netherlands: Elsevier.

Bonabeau, E. (2009). Decisions 2.0: The power of collective intelligence. *MIT Sloan Management Review, 50*(2), 45–52.

Carr, N. (2010). *The shallows: How the internet is changing the way we think, read and remember*. London, UK: Atlantic Books.

Chi, E. H. (2008). The social web: Research and opportunities. *Computer, 41*(9), 88–91.

Chi, E. H. (2009). Augmented social cognition: Using social web technology to enhance the ability of groups to remember, think, and reason. In 35th SIGMOD International Conference on Management of Data, Providence, Rhode Island, USA.

Chi, E. H., Pirolli, P., Suh, B., Kittur, A., Pendleton, B., Mytkowicz, T. (2008). Augmented social cognition. In AAAI Spring Symposium on Social Information Processing, Stanford, California, USA.

Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge, Massachusetts, USA: MIT Press.

Clark, A. (1999). Where brain, body, and world collide. *Cognitive Systems Research, 1*(1), 5–17.

Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive Science*. Oxford, UK: Oxford University Press.

Clark, A. (2007). Re-inventing ourselves: The plasticity of embodiment, sensing, and mind. *Journal of Medicine and Philosophy, 32*(3), 263–282.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York, New York, USA: Oxford University Press.

Clark, A. (2010). Memento's revenge: The extended mind, extended. In R. Menary (Ed.), *The extended mind*. Cambridge, Massachusetts, USA: MIT Press.

Clark, A. (2011). Finding the mind. *Philosophical Studies, 152*(3), 447–461.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19.

Clowes, R. (2015). Thinking in the cloud: The cognitive incorporation of cloud-based technology. *Philosophy & Technology, 28*(2), 261–296.

Clowes, R. W. (2013). The cognitive integration of e-memory. *Review of Philosophy and Psychology, 4*(1), 107–133.

Cooke, N. J., Gorman, J. C., Winner, J. L. (2007). Team cognition. In F. T. Durso, R. S. Nickerson, S. T. Dumais, S. Lewandowsky, T. J. Perfect (Eds.), *Handbook of applied cognition* (2nd edn). Chichester, UK: Wiley.

van Dijk, J. (2012). *The network society* (3rd ed.). London, UK: Sage Publications Ltd.

Donald, M. (1993). Precis of origins of the modern mind: Three stages in the evolution of culture and cognition. *Behavioral and Brain Sciences, 16*(4), 737–791.

Dourish, P. (2001). *Where the action is: The foundations of embodied interaction*. Cambridge, Massachusetts, USA: MIT Press.

Drain, C., & Strong, R. C. (2015). Situated mediation and technological reflexivity: Smartphones, extended memory, and limits of cognitive enhancement. In F. Scalambrino (Ed.), *Social epistemology and technology: Toward public self-awareness regarding technological mediation*. Lanham, Maryland, USA: Rowman & Littlefield International.

Fallis, D. (2011). Wikipistemology. In A. I. Goldman & D. Whitcomb (Eds.), *Social epistemology: Essential readings*. New York, USA: Oxford University Press.

Floridi, L. (2011). The construction of personal identities online. *Minds and Machines, 21*(4), 477–479.

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford, UK: Oxford University Press.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford, UK: Oxford University Press.

Gant, D., & Kiesler, S. (2001). Blurring the boundaries: Cell phones, mobility, and the line between work and personal life. In B. Brown, N. Green, & R. Harper (Eds.), *Wireless world: Social and interactional aspects of the mobile age*. London, UK: Springer.

Gemmell, J., & Bell, G. (2009). The e-memory revolution. *Library Journal, 134*(15), 20–23.

Gemmell, J., Bell, G., & Lueder, R. (2006). MyLifeBits: A personal database for everything. *Communications of the ACM, 49*(1), 88–95.

Giere, R. N. (2007). Distributed cognition without distributed knowing. *Social Epistemology, 21*(3), 313–320.

Giere, R. N. (2012). Scientific cognition: Human centered but not human bound. *Philosophical Explorations, 15*(2), 199–206.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, Massachusetts, USA: Harvard University Press.

Greengard, S. (2015). *The internet of things*. Cambridge, Massachusetts, USA: MIT Press.

Gregory, R. L. (1993). *Mind in science: A history of explanations in psychology and physics*. London, UK: Penguin Books Ltd.

Grier, D. A. (2013). Human computation and divided labour: The precursors of modern crowdsourcing. In P. Michelucci (Ed.), *Handbook of human computation*. New York, USA: Springer.

Halpin, H. (2013). Does the web extend the mind? In 5th Annual ACM Web Science Conference, Paris, France.

Halpin, H., Clark, A., Wheeler, M. (2010). Towards a philosophy of the web: Representation, enaction, collective intelligence. In Web Science Conference, Raleigh, North Carolina, USA.

Heersmink, R. (2015). Dimensions of integration in embedded and extended cognitive systems. *Phenomenology and the Cognitive Sciences, 14*(3), 577–598.

Heersmink, R. (2016). The cognitive integration of scientific instruments: Information, situated cognition and scientific practice. *Phenomenology and the Cognitive Sciences*, 15(4), 517-537.

Heintz, C. (2006). Web search engines and distributed assessment systems. *Pragmatics & Cognition, 14*(2), 387–409.

Helmstaedter, M., Briggman, K. L., Turaga, S. C., Jain, V., Seung, H. S., & Denk, W. (2013). Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature, 500* (7461), 168–174.

Hendler, J., & Berners-Lee, T. (2010). From the semantic web to social machines: A research challenge for AI on the world wide web. *Artificial Intelligence, 174,* 156–161.

Heylighen, F. (2013). From human computation to the global brain: The self-organization of distributed intelligence. In P. Michelucci (Ed.), *Handbook of human computation*. New York, USA: Springer.

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, Massachusetts, USA: MIT Press.

Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science, 2*(4), 705–715.

Hutchins, E. (2011). Enculturating the supersized mind. *Philosophical Studies, 152*(3), 437–446.

Hutchins, E. (2014). The cultural ecosystem of human cognition. *Philosophical Psychology, 27*(1), 34–49.

Janssen, J. H., Bailenson, J. N., Ijsselsteijn, W., & Westerink, J. H. (2010). Intimate heartbeats: Opportunities for affective communication technology. *IEEE Transactions on Affective Computing, 1*(2), 72–80.

Kanawattanachai, P., & Yoo, Y. (2002). Dynamic nature of trust in virtual teams. *The Journal of Strategic Information Systems, 11*(3), 187–213.

Kearns, M. (2012). Experiments in social computation. *Communications of the ACM, 55*(10), 56–67.

Khatib, F., Cooper, S., Tyka, M.D., Xu, K., Makedon, I., Popovic, Z., Baker, D., Foldit Players (2011). Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences*, *108*(47):18, 949–18, 953.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18,* 513–549.

Koriat, A., & Nussinson, R. (2009). Attributing study effort to data-driven and goal-driven effects: Implications for metacognitive judgments. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 35*(5), 1338–1343.

Kraut, R., Maher, M. L., Olson, J., Malone, T. W., Pirolli, P., & Thomas, J. C. (2010). Scientific foundations: A case for technology-mediated social-participation theory. *Computer, 43*(11), 22–28.

Kunze, K., Iwamura, M., Kise, K., Uchida, S., & Omachi, S. (2013). Activity recognition for the mind: Toward a cognitive "quantified self". *Computer, 46*(10), 105–108.

Landsness, E. C., Goldstein, M. R., Peterson, M. J., Tononi, G., & Benca, R. M. (2011). Antidepressant effects of selective slow wave sleep deprivation in major depression: A high-density EEG investigation. *Journal of Psychiatric Research, 45*(8), 1019–1026.

Law, E., & von Ahn, L. (2011). Human computation. *Synthesis Lectures on Artificial Intelligence and Machine Learning, 5*(3), 1–121.

Lintott, C. J., & Reed, J. (2013). Human computation in citizen science. In P. Michelucci (Ed.), *Handbook of human computation*. New York, USA: Springer.

Lintott, C. J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., et al. (2008). Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society, 389*(3), 1179–1189.

Ludwig, D. (2015). Extended cognition and the explosion of knowledge. *Philosophical Psychology, 28*(3), 355–368.

Lupton, D. (2013). Understanding the human machine. *IEEE Technology and Society Magazine, 32*(4), 25–30.

Lupton, D. (2015). *Digital sociology*. Abingdon, Oxon, UK: Routledge.

Lynch, M. P. (2014). Neuromedia, extended knowledge and understanding. *Philosophical Issues, 24*(1), 299–313.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., et al. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences, 97*(8), 4398–4403.

Malafouris, L. (2004). The cognitive basis of material engagement: Where brain, body and culture conflate. In E. DeMarrais, C. Gosden, & C. Renfrew (Eds.), *Rethinking materiality: The engagement of mind with the material world*. Cambridge, UK: The McDonald Institute for Archaeological Research.

Malafouris, L. (2013). *How things shape the mind: A theory of material engagement*. Cambridge, Massachusetts, USA: MIT Press.

Malone, T. W., Laubacher, R., & Dellarocas, C. (2010). The collective intelligence genome. *MIT Sloan Management Review, 51*(3), 21–31.

Marx, V. (2013). Neuroscience waves to the crowd. *Nature Methods, 10*(11), 1069–1074.

Mason, W. (2013). Collective search as human computation. In P. Michelucci (Ed.), *Handbook of human computation*. New York, USA: Springer.

Mason, W., & Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences, 109*(3), 764–769.

Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General, 137*(3), 422–433.

Matsumoto, T., Hashimoto, S., & Okude, N. (2008). The embodied web: Embodied web-services interaction with an umbrella for augmented city experiences. *Computer Animation and Virtual Worlds, 19*(1), 49–66.

Mayer-Schönberger, V. (2011). *Delete: The virtue of forgetting in the digital age*. Princeton, New Jersey, USA: Princeton University Press.

McBeath, M., Shaffer, D., & Kaiser, M. (1995). How baseball outfielders determine where to run to catch fly balls. *Science, 268*(5210), 569–573.

Menary, R. (2010). Dimensions of mind. *Phenomenology and the Cognitive Sciences, 9*(4), 561–578.

Michaelian, K., & Arango-Muñoz, S. (2014). Epistemic feelings, epistemic emotions: Review and introduction to the focus section. *Philosophical Inquiries, 2*(1), 97–122.

Michelucci, P. (Ed.). (2013). *Handbook of human computation*. New York, USA: Springer.

Michelucci, P. (2016). Human computation and convergence. In W. Bainbridge & M. Roco (Eds.), *Handbook of science and technology convergence*. New York, USA: Springer.

Miller, J. (2014). The fourth screen: Mediatization and the smartphone. *Mobile Media & Communication, 2*(2), 209–226.

Miller, M. (2015). *The internet of things: How smart TVs, smart cars, smart homes, and smart cities are changing the world*. Indianapolis, Indiana, USA: Que Publishing.

Nersessian, N. J. (2009). How do engineering scientists think? Model-based simulation in biomedical engineering research laboratories. *Topics in Cognitive Science, 1*(4), 730–757.

Neumann, M., & Cowley, S. J. (2013). Human agency and the resources of reason. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice*. London, UK: Springer.

Oksman, V., & Rautianen, P. (2003a). Extension of the hand: Children's and teenagers' relationship with the mobile phone in Finland. In L. Fortunati, J. Katz, & R. Riccini (Eds.), *Mediating the human body: Technology, communication and fashion*. Mahwah, New Jersey, USA: Lawrence Erlbaum.

Oksman, V., & Rautianen, P. (2003b). "Perhaps it is a body part": How the mobile phone became an organic part of the everyday lives of finnish children and teenagers. In J. Katz (Ed.), *Machines that become us: The social context of communication technology*. New Brunswick, New Jersey, USA: Transaction Publishers.

Parameswaran, M., & Whinston, A. B. (2007). Research issues in social computing. *Journal of the Association for Information Systems, 8*(6), 336–350.

Pfeifer, R., & Bongard, J. (2007). *How the body shapes the way we think: A new view of intelligence*. Cambridge, Massachusetts, USA: MIT Press.

Powell, A., Piccoli, G., & Ives, B. (2004). Virtual teams: A review of current literature and directions for future research. *The DATA BASE for Advances in Information Systems, 35*(1), 6–36.

Quinn, A., Bederson, B. (2011). Human computation: A survey and taxonomy of a growing field. In Annual Conference on Human Factors in Computing Systems (CHI'11), Vancouver, British Columbia, Canada.

Robbins, P., & Aydede, M. (2009). A short primer on situated cognition. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition*. Cambridge, UK: Cambridge University Press.

Rupert, R. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy, 101*(8), 389–428.

Salas, E., Fiore, S. M., & Letsky, M. P. (Eds.). (2011). *Theories of team cognition: Cross-disciplinary perspectives*. New York, New York, USA: Routledge.

Salerno, S., Zamagni, E., Urquizar, C., Salemme, R., Farn, A., & Frassinetti, F. (2012). Increases of corticospinal excitability in self-related processing. *European Journal of Neuroscience, 36*(5), 2716–2721.

Shapiro, L. (2007). The embodied cognition research programme. *Philosophy Compass, 2*(2), 338–346.

Shapiro, L. A. (2011). *Embodied cognition*. Abingdon, Oxon, UK: Routledge.

Sharma, V. K., Sharma, P., Deshmukh, R., & Singh, R. (2015). Age associated sleep loss: A trigger for Alzheimer's disease. *Bulletin of Clinical Psychopharmacology, 25*(1), 78–88.

Shrager, J., Billman, D., Convertino, G., Massar, J. P., & Pirolli, P. (2010). Soccer science and the Bayes community: Exploring the cognitive implications of modern scientific communication. *Topics in Cognitive Science, 2*(1), 53–72.

Simpson, T. W. (2012). Evaluating Google as an epistemic tool. *Metaphilosophy, 43*(4), 426–445.

Smart, P. R. (2012). The web-extended mind. *Metaphilosophy, 43*(4), 426–445.

Smart, P. R. (2013). Understanding the cognitive impact of emerging Web technologies: A research focus area for embodied, extended and distributed approaches to cognition. In 1st International Web for Wellbeing & Human Performance Workshop, Paris, France.

Smart, P. R. (2014). Embodiment, cognition and the world wide web. In L. A. Shapiro (Ed.), *The Routledge handbook of embodied cognition*. New York, USA: Routledge.

Smart, P. R. (in press). Emerging digital technologies: Implications for extended conceptions of cognition and knowledge. In: Carter, A. J., Clark, A., Kallestrup, J., Palermos, O. S., Pritchard, D. (Eds.), *Extended epistemology*. Oxford, UK: Oxford University Press.

Smart, P. R., Shadbolt, N. R. (in press). The world wide web. In: J. Chase, D. Coady (Eds.), *The Routledge handbook of applied epistemology*. New York, USA: Routledge.

Smart, P. R., Sycara, K. (2015). Situating cognition in the virtual world. In 6th International Conference on Applied Human Factors and Ergonomics, Las Vegas, Nevada, USA.

Smart, P. R., Huynh, T. D., Braines, D., Shadbolt, N. R. (2010a), Dynamic networks and distributed problem-solving. In Knowledge Systems for Coalition Operations (KSCO'10), Vancouver, British Columbia, Canada.

Smart, P. R., Huynh, T. D., Braines, D., Sycara, K., Shadbolt, N. R. (2010b) Collective cognition: Exploring the dynamics of belief propagation and collective problem solving in multi-agent systems. In 1st ITA Workshop on Network-Enabled Cognition: The Contribution of Social and Technological Networks to Human Cognition, Maryland, USA.

Smart, P. R., Simperl, E., & Shadbolt, N. R. (2014). A taxonomic framework for social machines. In D. Miorandi, V. Maltese, M. Rovatsos, A. Nijholt, & J. Stewart (Eds.), *Social collective intelligence: Combining the powers of humans and machines to build a smarter society*. Berlin, Germany: Springer.

Sparrow, B., Liu, J., & Wegner, D. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science, 333*(6043), 776–778.

Staley, D. J. (2014). *Brain, mind and internet: A deep history and future*. Basingstoke, England, UK: Palgrave Macmillan.

Sterelny, K. (2003). *Thought in a hostile world: The evolution of human cognition*. Oxford, UK: Blackwell Publishing.

Sterelny, K. (2010). Minds: Extended or scaffolded? *Phenomenology and the Cognitive Sciences, 9*(4), 465–481.

Stewart, J. R., Gapenne, O., & Di Paolo, E. A. (2010). *Enaction: Toward a new paradigm for cognitive science*. Cambridge, Massachusetts, USA: MIT Press.

Sutton, J. (2006). Distributed cognition: Domains and dimensions. *Pragmatics & Cognition, 14* (2), 235–247.

Sutton, J. (2010). Exograms and interdisciplinarity: History, the extended mind, and the civilizing process. In R. Menary (Ed.), *The extended mind*. Cambridge, Massachusetts, USA: MIT Press.

Sutton, J., Harris, C., Keil, P., & Barnier, A. (2010). The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences, 9* (4), 521–560.

Suzuki, K., Garfinkel, S. N., Critchley, H. D., & Seth, A. K. (2013). Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia, 51*(13), 2909–2917.

Swan, M. (2013). The quantified self: Fundamental disruption in big data science and biological discovery. *Big Data, 1*(2), 85–99.

Tasali, E., Leproult, R., Ehrmann, D. A., & Van Cauter, E. (2008). Slow-wave sleep and the risk of type 2 diabetes in humans. *Proceedings of the National Academy of Sciences, 105*(3), 1044–1049.

Theiner, G. (2014). Varieties of group cognition. In L. A. Shapiro (Ed.), *The Routledge handbook of embodied cognition*. New York, New York, USA: Routledge.

Theiner, G., Allen, C., & Goldstone, R. L. (2010). Recognizing group cognition. *Cognitive Systems Research, 11,* 378–395.

Toniolo, A., Ouyang, R. W., Dropps, T., Allen, J. A., Johnson, D. P., de Mel, G., Norman, T. J., Oren, N., Srivastava, M. (2014). CISpaces, towards a system for online collaborative argument construction and debate. In International Workshop on Arguing on the Web 2.0, Amsterdam, The Netherlands.

Tribble, E., & Sutton, J. (2011). Cognitive ecology as a framework for Shakespearean studies. *Shakespeare Studies, 39,* 94–103.

Valins, S. (1966). Cognitive effects of false heart-rate feedback. *Journal of Personality and Social Psychology, 4*(4), 400–408.

Vallar, G., & Ronchi, R. (2009). Somatoparaphrenia: A body delusion. A review of the neuropsychological literature. *Experimental Brain Research, 192*(3), 533–551.

Van Kleek, M., & O'Hara, K. (2014). The future of the social is personal: The potential of the personal data store. In D. Miorandi, V. Maltese, M. Rovatsos, A. Nijholt, & J. Stewart (Eds.), *Social collective intelligence: Combining the powers of humans and machines to build a smarter society*. Berlin, Germany: Springer.

Vincent, J., Haddon, L., & Hamill, L. (2005). The influence of mobile phone users on the design of 3G products and services. *The Journal of the Communications Network, 4*(4), 69–73.

Vygotsky, L. S., & Luria, A. R. (1994). Tool and symbol in child development. In R. Van Der Veer & J. Valsiner (Eds.), *The Vygotsky reader*. Cambridge, Massachusetts, USA: Basil Blackwell.

Webb, B. (1996). A cricket robot. *Scientific American, 275*(6), 94–99.

Weber, W., Rabaey, J. M., & Aarts, E. (2005). *Ambient intelligence*. Berlin, Germany: Springer.

Wegner, D. M. (1995). A computer network model of human transactive memory. *Social Cognition, 13*(3), 319–339.

Weiser, M. (1991). The computer for the 21st century. *Scientific American, 265*(3), 94–104.

Williamson, K., & Sutton, J. (2014). Embodied collaboration in small groups. In C. T. Wolfe (Ed.), *Brain theory: Essays in critical neurophilosophy*. Basingstoke, England, UK: Palgrave Macmillan.

Wilson, R. A., & Clark, A. (2009). Situated cognition: Letting nature take its course. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition*. Cambridge, UK: Cambridge University Press.

Won, A. S., Bailenson, J., Lee, J., & Lanier, J. (2015). Homuncular flexibility in virtual reality. *Journal of Computer-Mediated Communication, 20*(3), 241–259.

Woods, S. P., Iudicello, J. E., Moran, L. M., Carey, C. L., Dawson, M. S., & Grant, I. (2008). HIV-associated prospective memory impairment increases risk of dependence in everyday functioning. *Neuropsychology, 22*(1), 110–117.

Chapter 14
Models, Mathematics and Materials in Digital Architecture

Kåre Stokholm Poulsgaard and Lambros Malafouris

Abstract Today computing power and sophisticated digital tools are changing architectural design. Scripting and new design software is becoming ubiquitous and opens new opportunities for tech savvy designers, in turn computation seems set to fundamentally re-structure design practices. In the field of digital architecture, a number of critical voices have emerged from within its own ranks, vitally engaging with its theory and practice. These critiques universally assume that designers are, in fact, facing a new terrain for design thinking and that there consequently is a need to formulate a rationale for digital design research. If this is so, how can we begin to understand this new digital terrain, and what might its impact be on creativity and cognition? We approach these questions through the lens of Material Engagement Theory, exploring how computers and digital design research are changing the stakes for imaginative and creative thinking in architecture. We find that the potential of digital tools for bringing together vastly heterogeneous worlds might indeed extend the creative capacities of savvy designers, but that this relies on much more than a simple understanding of computation and involves materials, transactions and affect at several levels and temporal scales.

Creativity and Critique in Digital Architecture

Computing power and sophisticated digital tools are changing architectural design. Scripting and software offer unprecedented opportunities for tech savvy designers and hold the power to fundamentally re-structure design practices. As the field of digital architecture is establishing itself, a number of critical voices have emerged from within its own ranks, vitally engaging with its theory and practice

---

K.S. Poulsgaard (✉)
Institute for Science, Innovation and Society, University of Oxford,
64 Banbury Road, Oxford OX2 6PN, UK
e-mail: kare.poulsgaard@sjc.ox.ac.uk

L. Malafouris
Keble College and the Institute of Archeology, University of Oxford, Oxford, UK

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_14

(Terzidis 2006; Burry 2011b; Menges and Ahlquist 2011; Oxman and Oxman 2013; Davis 2013). There is a growing discussion of the conditions under which digital tools might set architects free to master hitherto unimaginable complexities, or whether, in worst-case scenarios, they degrade a century old craft and lead to cold and unimaginable digital reproduction. A basic assumption for these discussions is that designers are facing a new terrain in design thinking and that there is a need to formulate a rationale for digital design theory and practice (Oxman 2008, p. 102). If this is so, how can we begin to understand this novel terrain constituted by digital tools, and what might its impact be on creativity and cognition. There seems to be paradox at heart of the uptake of digital design tools: some designers, who master the subtleties of computational design thinking, are capable of highly original and open-ended engagement with digital extensions, while others primarily view Computer-Aided Design[1] (CAD) as a unimaginative tool for engineering efficiency. We will explore this seeming paradox through the lens of Material Engagement Theory, investigating how digital tools and design research is changing the stakes for imaginative and creative thinking. We proceed through two cases detailing specific moments in the development and application of computational design tools. These short case studies illustrate the ingenuity of three people working out tricky mathematical problems through design at different points in the 21st century. In Paris, the engineer Pierre Bézier drew on his experience as an engineer to create mathematically viable representations of curvature in CAD, and hence kick start the CAD/CAM revolution in design. In Barcelona, Antoni Gaudí worked through a constrained but sophisticated geometric language to create the iconic designs for the Sagrada Família; a century later architect Mark Burry and his team pioneered digital scripting and parametric modelling[2] to further develop these designs. At a first glance it seems the work of these designers proceeded in quite different ways. At a more fundamental level, however, the cognitive processes involved share a reliance on simplifying highly abstract numerical problems through manipulable models extending cognitive ecologies through paper, plaster or screens. These procedures show a number of cognitive mechanisms underlying the constitutive entwining of materials and mind at heart of material engagement theory; in turn, material engagement theory might help

---

[1] CAD/CAM, Computer-Aided Design and Computer-Aided Manufacture, encompass a growing suite of software used in digital design and manufacture. CAD often use 2D and 3D software environments for the design and documentation of digital products; CAM use software and programming to control production tools, including Computerised Numerical Control (CNC) milling machinery.

[2] Within digital design parametric modelling utilises the computer's ability to synthesise complexity in order to create models that describe explicitly defined relationships between terms in the model (cf. Davis 2013). This can perhaps best be described by an example: Rather than drawing a facade from the top down, so to speak, adding structure, walls, windows, ornament etc. you might instead describe formulas for the distance in-between windows and the size and shape of ornaments that might go there. Once form has been described in purely informational terms in the parametric model, its manipulation becomes much more fluid than anything drawing or traditional CAD can allow.

reveal cognitive processes underlying creative thought and practice at the human-computer interface.

A Material Engagement Approach to Digital Design

In his 2013 book, *How Things Shape The Mind*, Lambros Malafouris provides a comprehensive reframing of the problems of cognition and creativity by situating them within a relational framework of mind and material engagement. In this, thinking becomes a fundamentally situated activity where mind emerges through dynamic interaction between a variety of embodied, material and cultural domains in a specific environment. This situated and ecological perspective does away with Cartesian dualism by refusing to locate the origin of cognition and action either 'in the head' or 'in the world', rather it seeks to overcome these and other entrenched dualisms (e.g. mind/body, culture/nature) by showing how elements usually posited on either side of the pairs become inextricably entangled in ongoing processes of *mind in action*. Within this framework mind cannot be understood as anything but a relational process and we have to do away with notions of the brain as either originator or primary location for thought and action. Following this, we posit a locationally uncommitted theory of mind: neither bound by the head nor the body, mind emerges from interactional ecologies reaching well beyond the individual (Malafouris 2013a, p. 37). What matters is the coextensive entwining of cognitive processes, objects, materials and culture in situated and embodied action. Material Engagement Theory (MET) uses our deep phenomenological and creative engagement with material culture and our environments to explain how minds are enacted within relational fields. The change runs deep and privileging of mind over matter will not do within a relational framework where creative thought is inextricably tied up with action and mind is essentially co-extensive with tools and matter. In this perspective, form emerges through a morphological process owing as much to the intensive tensions and possibilities inherent in any material and tool as in design ideas held by the maker. The co-constitution of mind with matter leads to a *hylonoetic* ontology of thinking with and through matter[3] (Malafouris 2013b, p. 50)—embodied action brings forth the affordances of various material registers reshaping the enactive landscape of mind available to us. (cf. Gallagher 2014) This framework is processual, relational and intensive, and should provide us with new ways of thinking about digital design.

---

[3] From Greek *hyle:* matter, and *noetic:* relating to mind.

Three Theses: The Material Engagement Framework

Within this framework the key epistemological and practical problem becomes accounting for the ways in which different material and immaterial registers come to co-constitute and extend the plastic flexibility of minds in both situated action and long-term evolution. The answer reformulates and radicalises key elements from theories of extended and enactive mind with a strong materialist bent.

Material Engagement Theory champions a relational approach to thinking and doing: mind, action, form, all emerges through interplay of forces in a hylonoetic field extending beyond individuals into environment, culture, tools and materials. This perspective radicalizes extended mind theory by collapsing distinctions between internal and external elements of cognition. Mind becomes extensive rather than extended. In this we follow recent work by Hutto, Myin and Kirchhoff moving beyond theories of extended mind, by rejecting contents and representationrepresentation, while recasting the bounds of mind; deprived of content mind becomes enactive, essentially world involving, inherently extensive rather than occasionally extended (Hutto and Myin 2013; Hutto, Kirchhoff and Myin 2014). MET shares the anti-representationalism of radical enactivist theories of mind. Humans may be able to construct mental representations of almost anything, but we might not need these for many basic mental functions if we instead recognise that material engagement acts as a fundamental cognitive resource in its own right—that we think with and through things in action. What matters are neither representations nor their mental manipulation but rather the composition and coupling of forces making up the hylonoetic field. Analysing the problem of digital creativity within this framework requires attention to how problem spaces are configured and manipulated in design practice. And how this enactive and material engagement helps recast hard conceptual problems and support imaginative solutions.

MET proposes to move beyond mental representations by developing a concept of the material sign which opens for analysis of enactive signification. Material signs do not primarily embody a representational logic but an enactive one, they anchor and instantiates meaning through their very materiality; they stand for something, in ways that symbols cannot. Material signs are the actual physical forces that shape the social and cognitive universe, this depends on a strong commitment to enactivism and interactivity, where situated and embodied action brings forth the terms of interaction, such as design problems, tools, models and design ideas. As noted by Gallagher (2014) this transactional element of enactivism has strong roots in Dewey's pragmatism. Following in this vein, enactive signification highlights how creativity and cognition rides on movement, affect and transaction with shifting materials registers, and how this transaction profoundly shapes the perception and imagination of cognitive agents. This dynamical interactivity brings about the creative and flexible projection of mind into our material and symbolic surroundings.

This commitment to interactivity has consequences for concepts of agency. In the hylonoetic perspective action does not primarily arise from mental plans and representations, but rather within the complex interplay of forces and affordances in a given relational ecology. In this view agency as an emergent property can neither be reduced to human nor non-human components. Agency only makes sense in an interactive, relational domain where all elements have had their power to deeply influence each other restored. The notion of material agency seeks to capture this domain of possibilities and the fluid and shifting properties of (extensive) mind in action across these. When analysing processes of digital design, we can neither reduce our explanations to the creative genius of the individual designers nor to the cold hard logics of computational tools; rather we should see how intent, experience and logics flow together in situated social and material interaction.

MET thus adds a unique temporal perspective to current debates in the philosophy of mind; at heart MET is concerned with the intimate link between being and becoming, seeking to develop a framework for apprehending this question across a variety of domains and time-scales. This commitment to processes unfolding at various timescales from the short to the long term presents a relevant and dynamic framework for interrogating contemporary design research within a broader context. Attuned to evolutionary and historical context and contemporary qualitative change, MET opens for a philosophically tinged anthropology of the current digital and informational revolution (discussed in the final section). We humans have plastic minds enfolded within plastic culture, and constantly alter our cognitive ecologies. By interacting with tools and materials we change the projective flexibility and material make-up of our minds (cf. Kirsh 2009; Hutchins 2010). Within MET this is called metaplasticity. The main thrust of this is a long scale view recognising that the crux of human cognitive evolution: "is not to be found in the ever-increasing sophistication or specialization of a modular mind, but in an ever-increasing projective flexibility that allows for environmentally and culturally derived changes in the structure and the functional architecture of the brain's circuitry." (Malafouris 2013a, p. 46).

The Times of Digital Design Research

The relational and flat ontology at the heart of MET makes non-anthropocentric analyses of computation and design possible by advancing a radically different view of thinking and doing from representationalist theories of mind. This framework allows for technical and informational logics informing analysis of creative work as it entails a shift in perspective towards an interactive and ecological analysis. Investigating digital design as a creative co-evolution of tools, practices and minds lets us discover at least two distinct but mutually supportive timescales at play in contemporary practice: one long term, having to do with the development of digital geometry and sophisticated mathematics; the other short term and affective, tied to the individual designer's history of engagement with tools and software. In both dynamic interaction with tools and materials play evident and surprising roles.

Designing Curvature for Numerical Control at Renault in the 1960s

Non-Uniform Rational B-Splines[4] (or NURBS for short) is the mathematical model for generating and manipulating curves and surfaces in most 3D and CAD tools today. Their graphical representation as curvature with a series of floating control points connected to knots on the surface, makes their manipulation easy and intuitive, far removed from the complexities of the underlying polynomial expressions that make them work mathematically. Their robust underlying mathematics and their intuitive control means that they offer great flexibility and precision for both mathematically defined and modelled curves and surfaces and they quickly became an industry standard after their development in the 1970s. Today they are the lingua franca of digital design: “Want to talk business, learn to talk NURBS.” (Piegl and Tiller 1997, p. 9). But how did we learn to talk NURBS in the first place? To better understand cognition and creativity in digital design it might be worthwhile to analyse the early development of NURBS and their contemporary use.

CAD and NURBS have their roots in 1960s Paris where the futuristic designs of Citroën and Renault lead to the almost simultaneous development of computational curvature.[5] Around this time automobile design progressed, hand and craft, through a number of materials and models until a single full-scale template was created in wood and clay. This one model would serve as the master for all cars produced. The process was slow and as more and more of the production became automated via computerised numerical control (CNC) machines, the design and production teams needed simple ways of transferring the curvature of models to accurate numerical descriptions for production. Bézier (1983, 1986, 1993, 1998) describes the remarkable development of Bézier curves and surfaces at Renault, the first published solution to this problem and an important predecessor to contemporary NURBS. Bézier was not a mathematician but an engineer, the son and grandson of engineers, and emphasize that early CAD at Renault was developed by engineers and workmen: “not bad at math” but “far from ‘chartered’ mathematicians.” (Bézier 1998, p. 39) This led to a heterodox and highly original path towards a solution to the non-trivial mathematical problem of describing freeform curves and surfaces. Bézier relied on a series of conceptual transformations of the problem, sketched out

---

[4] Traditionally a spline was a long piece of pliable wood or metal whose material composition made it form smooth curves when bended and held in place by a series of ‘ducks’; this allowed for optimisation of curved forms and was particularly valued in shipbuilding and design where their use can be traced back to the early AD Romans. Splines were first described mathematically in 1946, but their material forebears have continued to inspire mathematicians and engineers developing curvature for digital design tools (Farin 2002).

[5] Computerised curvature was first developed by Paul De Casteljau (in 1959), a young mathematician at Citroën who were not allowed to publish his results, and subsequently by Pierre Bézier (in the early 1960s), an engineer at Renault. While arriving at similar solutions, they proceeded along quite different paths.

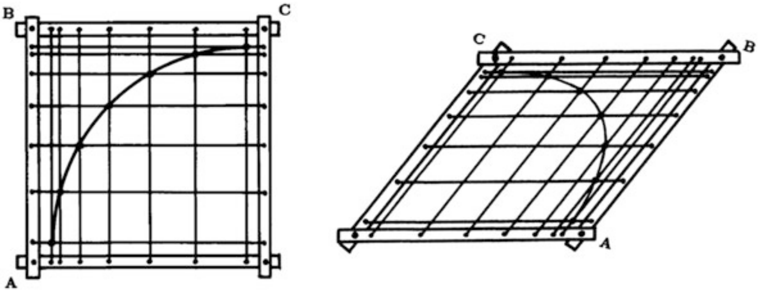


Fig. 14.1 Bézier's sketches. Imagining a flexible rectangular grid, Bézier was able to transform a curve using only three control points (Bézier 1993 © Elsevier)

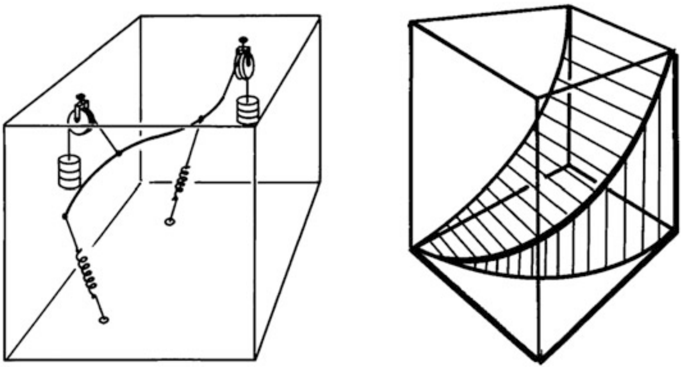


Fig. 14.2 Bézier's sketches. Moving from 2D to 3D space, Bézier tried out a number of ideas before realising that he could describe a transformable curve inside a cube (Bézier 1993 © Elsevier)

initially with pen and paper and since tested using a flatbed plotter and CNC milling machines. While tasked with finding a way to transform curvature from models into accurate numerical descriptions, he realised early on that this had the potential to become a two-way conversation if he also managed to create a simplified interface for manipulating numerical curves in an intuitive way (Bézier 1986).

Bezier's first conceptual draft was a simple square frame containing a grid, whose intersecting strings describe control points on a given curve. By transforming the square frame into a diamond the intersecting control points would change and distort the curve (Fig. 14.1).

This simple displacement fitted well with underlying mathematics but was too restrictive, it did not allow for freeform curves in 3D space, and a more flexible solution was required. Bézier sketched out a number of spatial contraptions, at one point mulling over a room with a hanging sweep whose curvature was determined by a number of springs, pulleys and weights. While this would allow for freeform curvature it was highly impractical. But, retaining the 3D perspective, Bézier came up with a simple and elegant solution (Fig. 14.2).

Returning to the idea of a frame, he realised that 3D parametric curves could be defined by the intersection of two cylinders within a cube. Once formulaically described within this, the curvature could be distorted and calculated by transforming the outer cube and keeping track of three of its edges. This allowed for a simple way to describe and transform free-form curves, while the underlying math could rely on robust and well-known polynomial expressions. While these consecutive steps might seem unnecessary or fanciful once the solution is presented, Bézier is adamant about their importance: "the various steps of this conception have a point in common: Each idea must be related to the principle of a material system, simple and primitive though it may look, on which a variable solution could be based." (Bézier 1993, p. 9). In this way, solutions rely on the material transformation of the problem using pen and paper and come to echo the engineer's affective memory of working with frames, splines and tools. In the end, Bézier's solution would have a tremendous impact, not just on CAD but also on computer graphics, fonts and design, due to it being simultaneously sophisticated and intuitive. With this interface between form and mathematics new inferences became both possible and in many cases almost automatic. For some it seemed not much had gone into their innovation in the first place; Bézier remembers an unimpressed executive at Renault telling him that if his system was that good: "the Americans would have invented it first." (Rogers 2000, p. 36).

Digitising Plaster Models at the Sagrada Família Church in the 1990s

Architect Mark Burry is among the second generation of pioneers in digital design and has consistently been pushing the limits of practice through imaginative and creative work with digital tools and methods. Since the 1990s he has overseen the completion of Antoni Gaudí's Sagrada Família basilica in Barcelona, and through this work has championed an open-minded and productive integration of digital and traditional techniques. While his trajectory might not be typical for digital designers today, Burry's work reveals how we might begin to apply aspects of material engagement theory for understanding contemporary digital design and imagination.

In a "cultural defence" of digital tools and scripting for design, Mark Burry recalls his first experiences with CAD, a decade after we left off with Paul Bézier in the 1960s: "once a week for a semester we went down to the university's computer centre where we would do the maths and valiantly punch cards to input the computer with the necessary analogue binary data. At the end of the semester I had computer-drafted a cube that could be viewed in perspective; an inauspicious introduction to design computation." (Burry 2011b, p. 14) Cambridge was at the absolute forefront of digital design and computer graphics in the 1970s but this did little to advance Burry's enthusiasm. This changed however, in the late 1980s, when he was doing research as a student on Antoni Gaudí's Sagrada Família church in

Barcelona. The Sagrada Família is a monumental project in every way. Begun in 1882 and set to be completed under Burry's supervision in 2026, its design and construction will span more than 150 years and combines techniques from Gaudí's early Gothic revival through his later use of sophisticated geometry, to computational research and rapid prototyping today. Gaudí revised the plans for the church numerous times and devoted the last 12 years of his life to developing a systematic design for the unfinished parts of the building. While he is known for his frivolous freeform designs there is a remarkable and sophisticated use of geometry at play in this late stage, which was developed largely through 1:25 and 1:10 scaled plaster models. As Gaudí never wrote down a word, these models are crucial 3D evidence of his transition from freeform design to advanced use of geometry (Burry 2015, p. 6). While it is beyond the scope of this paper to dive into the impetus and implications of these geometrical experiments, they are worth visiting in brief to give an idea of the challenge faced by the teams working to complete the Sagrada Família. For a start, most of Gaudí's plaster models were trashed during the Fascist reign in the decades after his accidental death in 1926. The models have since been painstakingly dug out, documented and restored. The construction of the Sagrada Família is supported by an on-site team of master modellers who can trace an unbroken lineage back to craftsmen active at Gaudí's time. Engaging with these and working with Gaudí's successors, Burry has built up what is most likely an unrivalled understanding of Gaudí's geometric thinking today (Burry 1993, 2005, 2011a, 2011b). Computational analysis and experimentation have played a key role in this, initially tasked with extracting and describing geometric parameters for restored models, Burry realised the potential of scripting for design, using digital tools to map and explore the geometries inherent in Gaudí's system. Gaudí's late design for the Sagrada Família came to rely almost exclusively on the geometry of doubly ruled surfaces. These are warped surfaces that can be described by an array of straight lines lying across the surfaces in two directions; they combine visual sophistication with geometric simplicity. There are only three doubly ruled surfaces: the hyperbolic paraboloid (saddle shape), the hyperboloid of revolution of one sheet (hourglass shaped) and the helicoid (spiralling surface), and Gaudi used these to structural and visually stunning effect (Burry 2011a). While turning knowledge of this geometry into blueprints for the construction of Gaudí's clerestory window, Burry first realised the enormous potential of digital scripting for design.

The clerestory windows of the Sagrada Família church run along the length of the central nave 35 m above the aisle below. They are an important part of the full design but their geometry remained to be extracted from experimental models left unfinished by the time of Gaudí's death; the Sagrada Família team faced a difficult task in developing highly accurate documentation for their production. In Burry's own words: "This work took me from being deeply embedded in traditional media through to developing into a total believer in scripting by the time the window was built." (Burry 2011b, p. 127). The clerestory window is made up of three intersecting hyperboloids of revolution. Two round below a third elliptical to create a complex and beautiful 3D geometry of curvature and ridges derived from the intersecting of these forms. Finding the optimal shape and intersection of these is a non trivial problem: a hyperboloid of revolution has nine variables influencing its situation with respect to its neighbours: three coordinates of location, three axes of rotation, and three constants determining width and height of the opening and the degree of the slope (Burry 2011b, p. 134). Even a slight change in one variable can have a large effect on its appearance, which exaggerates further when the forms intersect as in the clerestory window. Taken together the intersection of these nine variables produces an almost endless variability and, one can imagine, a headache of similar magnitude when trying to find the optimal solution. For Gaudí, the answer lay in iterative material engagement, sketching the shapes before having his plaster makers create the form and working through their intersection manually (Burry 2011b, pp. 131–133). This eliminated all abstract calculations, but was time consuming and inexact; the scaled models show clear signs of rasping where these did not intersect smoothly. Seeking to develop exact and precise documentation for the production of the clerestory window, Burry enlisted the help of a computer scientist who developed a genetic algorithm to compute the data for compiled points and generate the fit that matched best for all nine parameters. This could then be exported to parametric software that could generate a 3D model (using NURBS curvature and Boolean integration of the hyperboloids) (Fig. 14.3).

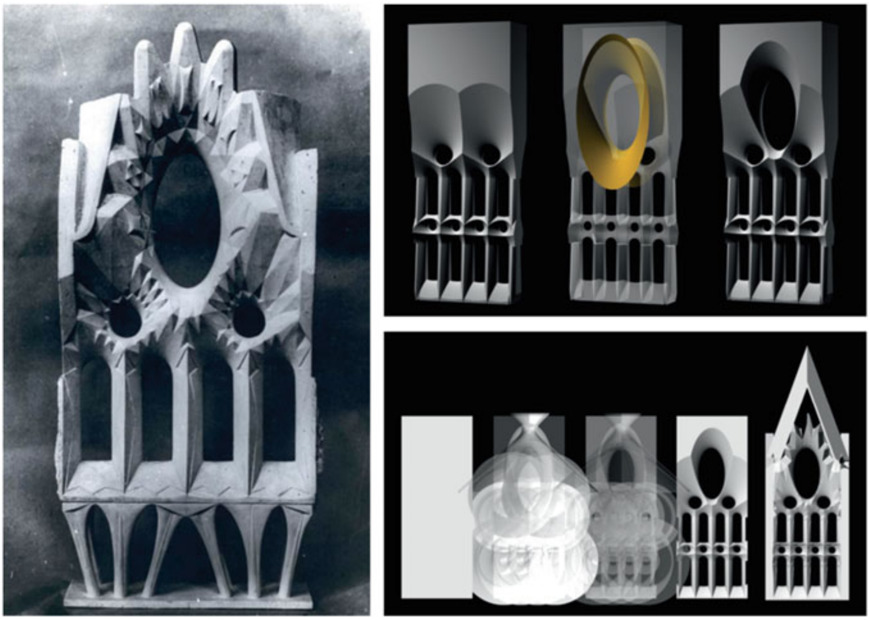


Fig. 14.3 Sculpting Gaudí's clerestory window. *To the left* a photo of one of Gaudí's original 1:10 scaled models made shortly before his death in 1926 (© Temple of the Sagrada Família); *to the right*, making the digital 3D model by removing hyperboloids using Boolean subtraction (Burry 2011b © Mark Burry)

That was not the end of it however. Running the algorithm required a lot of fiddling with the measured input variables, and even changes of half a degree along any axis had visually disproportionate effects. The 3D models continuously had to be adapted by eye to find as precise as possible a match with Gaudí's originals (Burry 2011b, p. 143). This blend of traditional methods, abstract scripting, 3D models and ongoing visual tweaking, shows how rich a field the digital design research on Gaudí's clerestory window relied on. The team proceeded from a pre-given design object, but explored this via the variability and subtlety available through digital design tools. While this is a lot more advanced than traditional CAD would allow for, it is in some respects a conceptual continuation of the work of Bézier, in that it relies on integrating a deep understanding of morphology with the accuracy of mathematics through manipulable 3D representations of geometric form. As such the digital design is at once conceptual and materially anchored, made available for more intuitively embodied experience through ongoing manipulation and visual judgement. The work of Bézier shows how conceptual transformation of problems anchored in material aids and tools can lead to new and imaginative solutions; Gaudi and Burry further underscores how this process relies on a continuous restructuring of the plastic hylonoetic field supporting creative thought and design.

Enactive Signification and Computational Design Thinking

These short case studies illustrate the ingenuity of three people working out difficult mathematical problems through design at different points in the 20th and 21st century. Their practices show a number of cognitive mechanisms underlying the constitutive entwining of materials and mind at heart of material engagement theory, specifically: enactive signification; materially anchored conceptual blending; and the ongoing co-constitution of tools, practices and institutions at different temporal scales. All mechanisms have salience for understanding digital cognition and creativity.

Conceptual Blending and the Development of Curves in CAD

Conceptual blending theory (Fauconnier 1997; Fauconnier and Turner 1998, 2002) refers to non-metaphoric mappings in the conceptual system of a cognitive agent. In this, structure from two or more input spaces is selectively projected into a separate, blended space, which thereby develops emergent cognitive properties not available from the separate inputs. It is routine activity and while Fauconnier and Turner operate in mental realms, Hutchins (2005) has shown how blends often rely on material anchoring, that is externalisation of the blend in materials and tools that provide stability while removing cognitive load. Indeed, material anchoring greatly enhances the emergent possibilities available to cognitive agents by introducing interactional ecologies for manipulation and tinkering (Vallée-Tourangeau and Villejoubert 2013). Materially anchored conceptual blending is the crux of enactive signification and the constitutive entwining of cognition with the material world; in light of the above we can begin to see how this might work in the digital realm.

In blending theory, two or more input spaces contribute structure selectively via projection to a third space, the blend, which as a result develops a new emergent structure. In Bézier's case we see how structure from abstract mathematics is projected together with a variety of conceptual domains to create the Bézier cube, a blend selectively incorporating elements from polynomial expressions, Cartesian geometry and algebraic functions, but also from more immediately material domains such as wooden frames and the deformation of metallic splines under the give and take of spring-loaded mechanisms (Bézier 1993). Through successive blends the problem is transformed, first through a simple deformable grid, then through the spatial domain of a spring loaded room; these subsequently blend to form the final space of the Bézier cube, mixing the transformations of the 2D frame with the Cartesian space of the 3D room. While it might be tempting to think of this as a purely mental blend, Bézier himself highlights the role played by material tools and knowledge from model making. Besides pen and paper, he drew intuitively on the experience and craft of engineers and workmen for developing his curves, and from the beginning the development was anchored materially in technical tools. Hutchins describes material anchors for conceptual blends this way: "A mental space is blended with a material structure that is sufficiently immutable to hold the conceptual relationships fixed while other operations are performed." (Hutchins 2005, P. 1562) The development of Bézier curves and their subsequent use was performed with the help of a large flatbed plotter, allowing the team to experiment, review and revise their work by transforming mathematical expressions into visual curvature (Bézier 1998, p. 40). In fact this flatbed plotter was such an integral element of the functioning of Bézier curves, that CAD pioneers visiting Renault in the 1960s vividly remember "the crazy way Renault design surfaces" and the "new system on the graphics output of the time, a large flatbed plotter." (Forrest and Riesenfeld in Rogers 2000, p. 13, 38). How might this plotter work as a material anchor for the conceptual blends at heart of Bézier curvature? Hutchins goes on to explain that the crux of material anchoring lies in the functional coupling between material and conceptual domains: "If conceptual elements are mapped onto a material pattern in such a way that the perceived relationships among the material elements are taken as proxies (consciously or unconsciously) for relationships among conceptual elements, then the material pattern is acting as a material anchor." (Hutchins 2005, p. 1562) While the plotter itself might not constitute such a pattern, its output—curves on paper—certainly does and actively comes to stand for the relationships among the underlying polynomial expressions and specific variables. Creating simple geometric proxies for the exact numerical representation required by CNC machinery was the explicit motivation for Bézier's work, and further to do it in such a way that designers and drafters could easily experiment within the interface environment without having to worry about the underlying numerical multitudes. In Bézier's (1986, p. 156) own words: "Since it is vital to give stylists or designers the possibility to see and touch without delay a 3-D rendering of the object they have conceived, it is necessary to provide them with machines able to carve very rapidly a large part of a car or, even better, a complete one." In this way the blend becomes stabilised, transferable and transmutable, relying on the material and cultural anchoring springing up around CAD and CNC milling technology in the 1960s and beyond.

Conceptual blends evolve through *composition*, *completion* and *elaboration* (Fauconnier 1997, pp. 150–151). In composition selective inputs blend to make new relations available that did not exist in the inputs (e.g. intuitively graspable curves with high numerical accuracy). Completion of the blend in this way merges knowledge of background frames, cultural and material models into a materially anchored composite, the blend, with its own self-contained structure (e.g. the Bezier system anchored in plotter and CNC milling technology). This is a highly interactive process. The blend is established and elaborated through active work performed within the blend according to its own emergent logic. We see this elaboration at play in the practice of mathematicians and computer scientists working from Bézier curves to develop the NURBS at heart of 3D design today (see Farin 2002). In computer science and maths then, we might expect conceptual blending to take on a recursive, almost fractal nature, as completed blends act as input spaces for yet other blends and so on (Fauconnier and Turner 2002; cf. Lakoff and Núñez 2000; Núñez 2010). In developing interfaces for computers and digital tools, abstract knowledge gains material anchoring which gives it an intuitively graspable dimension allowing non-experts to work within the completed blend by relying on bodily grounded experience such as spatial logics, visual perception and haptic, material engagement. This goes some way in explaining the long term cognitive and cultural development underlying contemporary design research, it does not however, explain how designers engage creatively with these ecologies to come up with original and surprising solutions.

Enactive Signification and the Aesthetics of Practice

The application of blending theory becomes a bit more muddled when turning to the story of design research and construction on Gaudí's clerestory window. As with Bézier's curves, the abstract mathematical problem faced by Burry and his team—finding the optimal integration of three hyperboloids while respecting Gaudí's unfinished experiments—had to be recast through material and digital tools to allow for new ways of manipulating variables while evaluating their output. Developing the parametric 3D model for this required input from restored plaster models (form, measurements), the geometry of doubly ruled surfaces (the formula for hyperboloids of revolution), computer science (a genetic algorithm) and 3D graphics and operations (NURBS and Boolean logic). We might forward the 3D model as a conceptual blend of these areas, but if so, it is not entirely clear, what is selectively blended and what is left out, as it is rather a case of all these dimensions being integrated into the parametric model along a historical dimension (cf. Burry 2011b, p. 94). This happens in progressive steps, non-crucial information might be left out from step to step, but is retained in the explicit history of the final model. Burry explains (Burry 2011b, pp. 137–143) how measurements from remade plaster models were used to generate a variety of digital 2D curves, up to a hundred at a time due to their great variability, which were transformed to digital 3D models of hyperboloids of revolution. Searching for the optimal intersection of these was done using a genetic algorithm. Again, due to high variability, the algorithm was run numerous times over several weeks and inputs were continuously evaluated and tweaked. The algorithmic output was a script for automatically generating a parametric 3D model of the clerestory window using the geometric primitives of CAD software. At this point the hundreds of iterations from earlier steps are missing from the final 3D model, but their general principles remain in the selected geometry and versions continued to play a part through ongoing optimisation. The best fit for the intersecting hyperboloids of the window was found by tweaking forms within the parametric 3D environment while keeping an eye on Gaudí's restored plaster models, but this is only the tip of the iceberg of an evolving process incorporating knowledge from plaster makers, mathematicians, computer scientists and architects. With the explicit history and clearly defined steps at heart of digital design research, processes remains open—at least in principle—for feedback, experimentation and the production of new knowledge along the way. This layered sequence works by allowing designers selective access to data and representations so that relevant or surprising combinations might be made visible at any one time (cf. Cross 2006, p. 37). Completion and elaboration takes on importance in this domain, as the establishment of a workable parametric model relies on successfully developing a sequence of steps with a self-contained logic within which design problems can be processed and transformed in order to arrive at a satisfying solution; in this case a mathematically precise model with a visually optimal fit with Gaudí's unfinished plaster models. While completion in Fauconnier and Turner (1998, p. 144), relies on implicitly filling in the missing pieces from background knowledge and frames, the digital model is differently explicit, containing all the sequential steps, along with the inputs and outputs of each. Not all of these are made visible at any one time however. Burry explains how experimentation at each step became key for the successful construction of the clerestory window, and how the emergent structure of the final window was progressively refined through material engagement with plaster models, design scripting and visual optimisation. The process brings together embodied experience of engaging with a range of material registers, and the various steps are materially anchored in the models, screens and interfaces allowing for manipulation of the problem space. It seems the case of sophisticated digital design research is susceptible to blending theory with the added element that what goes on is not necessarily a one off integration, as much as an explicit historical layering, allowing for constant realignment, shifts and integrations between layers. The success of the resulting parametric model lies in its possibilities for functional simplification, that is the way it allows designers to selectively zoom in on, mix and

Fig. 14.4 Finding points of intersection on Gaudí's clerestory window. *To the left*, plaster makers using traditional techniques (Burry 2011b © Mark Burry); *to the right*, overlaying the digitised plaster models with variable hyperboloids of revolution allowed the team to find the best fit (Burry 2011b © Mark Burry)

process specific aspects while ignoring others. This privileged view permits research into dynamic and complex systems at a human scale. In the case of Gaudí's clerestory window, Burry and his team progressed towards a parametric 3D model allowing them to manipulate their findings on screen to visually find the best fit between digital and material models. In this, the visual model freed the mind of the designers from keeping a vast array of numbers in their heads and rather concentrate on the aesthetics of their geometric integration. Working through the problem using the affordances of various material and immaterial registers remained important throughout this process and provisional material modelling using plasticine was undertaken to "tune" the parametric scripting (Burry 2011b, p. 132) (Fig. 14.4).

In a recent article, Burry (2015) takes stock of the different practices and material registers at play in the ongoing construction of the Sagrada Família cathedral; these range from the century old craft of model makers to digital scripting and rapid prototyping techniques and integrate onsite as continuous rather than disjointed fields. While this breadth of tools and techniques might be unique in architecture today, the insights it holds for understanding digital practice should be of broad interest. There remains much to be explored to develop this into a fully-fledged theory of digital design, but analysing the construction of Gaudí's rose window through the lens of material engagement theory reveals the inherent multi-layered nature of digital design research, and the richness of its practice. This seems to rely on recursive conceptual blending, as well as material and conceptual anchors (computers and models) that turn numerical data into meaningful aesthetic experience. What constitutes meaningful experience in this domain is not a given but rather relies on the history of the individual designer. Practice attunes attention; it involves learning techniques while developing an understanding of tools, logics and morphology not to mention their purposeful combination (Menges and Ahlquist 2011; cf. Lave and Wenger 1991). "Such oscillations unfold within an aesthetics of practice which derive from the sensory and emotional relationships between the forms of things and the appreciations of the human body." (Gosden and Malafouris 2015, p. 707; cf. Malafouris 2011, 2013b and 2014) In digital design research the computer becomes a constitutive element in a larger hylonoetic field spanning informational, conceptual and material realms. The many registers are integrated conceptually through embodied interaction with design tools, combining a deep understanding of computation with a feel for materials and morphology. This tight coupling between a variety of material and conceptual fields alongside personal experience and affect is the crux of enactive signification in the digital realm; enactive signification in turn allows for what we might call "designerly imagination", aha moments when disjointed tools and experiences flow together and new opportunities emerge (cf. Cross 1982; Archer 1995; Frayling 1994).

Enactivism and Creative Digital Design Thinking

Architects Achim Menges and Sean Ahlquist describe computational design thinking like this: "Fundamental to computational design is the understanding of how systems, as form and as mathematical ordering constructs, operate. Fundamental to understanding their operation is the level of prediction contained in the model which envisions the system." (Menges and Ahlquist 2011, p. 16) In this, the position of the designer becomes conspicuous, explains Menges and Ahlquist, as computation relies on the purposeful establishment and execution of rules for the development of forms, but *not* explicit descriptions of final forms themselves. This shift from finished object to process, from free form design to rule-based exploration, is at the heart of digital design (cf. Cross 2006, p. 32; Simon 1996, p. 124) but within this definition it remains unclear exactly how imagination and creativity emerge from digital design research. Running a sophisticated model or script can reveal new insights about a complex problem; it is less clear how new opportunities and insights emerge in the process of establishing the model itself but enactive signification might help explain this. If process, variability and emergence are cardinal points of computational design thinking, the experience and knowledge of the individual designer seem vital to the intentional but creative navigation of complex design environments today. When constructing a model, a design team needs to project and blend insights from software, mathematics, materials and morphology onto a variety of materially anchored digital ecologies, selectively integrating different layers of information and logic to establish the space within which to explore solutions. In cognitive theory, David Kirsh (2009) has compared cognitive projections to a way of augmenting reality. In his own words cognitive projections describe: “a way of ‘seeing’ something extra in the thing present (…) a way of *augmenting* the observed thing, of projecting onto it” (Kirsh 2009, p. 2310, emphasis in original). Things serve as anchors for projection, which in turn goes beyond what is immediately present for material and symbolic engagement by calling up additional possibilities (Hutchins 2005; Kirsh 2013). Kirsh primarily deploys this line of thinking within a representationalist theory of mind, but as he himself notes, there is no reason that this has to be so, and in fact notions of projection sit well within enactive frameworks (Kirsh 2009, p. 2314). In some enactive theories, everyday perception contains imaginative elements, ways of ‘seeing the future’ as Kirsh puts it, for example in the way perception and imagination flow together when we perceive an object from one angle and automatically envision the back of it (O’Reagan and Noë 2001). In contemporary enactive theories of mind, this imaginative engagement with objects relies on embodied interaction rather than pre-conceived notions in the form of representations (see, Hutto 2015). This adds further dimensions to the richness of experience and aesthetics of practice forwarded above by explaining how creative opportunities reveal themselves through rich material and symbolic engagement. Malafouris proposes cognitive projection and conceptual blending as the basic non-representational mechanisms for the “dense structural coupling” between mind and matter that is at heart of enactive signification; the mechanisms through which the material sign brings forth the world (Malafouris 2013a, p. 100). Through this structural coupling, the aesthetics and affects of embodied practice come to shape perception and imagination, and the creative enactment of digital design research. Menges and Ahlquist point to an “understanding of how systems, as form and as mathematical ordering constructs, operate” as the crux of this process; Burry has shown that this exceeds mental or static representations and rather develops through creative engagement with materials and tools as well as software and logics of abstraction and concretisation (morphology and form). In this way computational design thinking proceeds not only through rule-based exploration but also through creative combinatorics, sense-saturated interactivity that reveal new opportunities in the complex and dynamic systems that designers brings together through material engagement (cf. Steffensen 2013).

The inherent historical coupling between cognitive agent, tools and environment is essential to the enactive view of mind at heart of material engagement theory (Gosden and Malafouris 2015). Even for highly abstract mathematical entities, such as polynomial expressions or doubly ruled surfaces, a wide variety of material and conceptual registers and anchors are recruited to allow for their purposeful manipulation in the realm of aesthetic experience. It is this transactional engagement with shifting material social and informational registers within the hylonoetic field that gives rise to original insights and imaginative ideas (cf. Gallagher 2014, p. 116). For experienced digital designers, computers and digital models help support creative integration between diverse material and immaterial registers, and this integration becomes salient through the affect of personal experience. Creativity arises from the many possibilities for combinations and permutations within this field, and sense-making through embodied practice becomes a key activity for exploring the affordances of new combinations. This combinatorial game only expands with the richness of personal experience. Digital design research becomes vital for the individual designer, not necessarily for producing specific results, but for deepening the experiential field from which creative and meaningful material engagement with digital tools can emerge and evolve. Design research is a mode of materially and bodily anchored thought in action, and becomes recursive—for both individual designers and the design community at large—in the way it is vitally engaged with extending the possibilities for this as a cognitive process.

Metaplasticity and the Intersecting Times of Digital Design

The recursive nature of design research as a cognitive and creative process is active at several temporal scales. From the minutia of absent minded fiddling with tools, through the conscious and unconscious intuitions borne from decades of affective experience, the entwining of mind with material and symbolic realms shape creative practice; as historical cultural and cognitive evolution underpins our development with and through tools, there is a long-term influence active as well (Malafouris Malafouris 2013b, 2015). We humans are the only species with plastic minds enfolded within a plastic culture, and this metaplasticity opens for a rich field of creative and imaginative engagement with our surroundings. This engagement is contingent on both immediate and past processes of mind, and in turn shapes possibilities for future development. Analytically it might make sense to keep timescales separate, but we could also enquire into how they run together through situated and embodied practice. By analysing two instances of creative digital engagement—one an implicit prerequisite for the other—we have drawn out a single example of how this might be done to understand digital design in architecture. Of course a few generations hardly qualify as longue durée, but the genealogy of digital curvature could be extended much further to encompass the development of curves and splines when technical drawing was introduced to boatbuilding in the 1600s (Johnston 1994); the practical geometry of gothic arcs (Turnbull 1993), or even further back, to early use of drawings, tools and templates amongst Greek and Roman builders (Senseney 2011). At each instance changes in our extended cognitive ecology seems to have reshaped the mind of our species and in turn recast our possibilities for engaging creatively with the world.

Our creative lives—thinking, doing and making—unfolds with and under the pressure of heterogeneous forces coming together to afford and constrain practical and cognitive opportunities available to us. Hutchins (2010) has used the notion of cognitive ecologies to explain the dynamic interaction and mutual co-constitution of materials, tools and culture. Cognitive ecologies contain past and present tools and techniques and become stabilised though material and cultural anchors. Cultural knowledge and cognitive evolution are not intracranial processes; they are, rather, infused and diffused into settings of practical activity; they are constituted by experience within these settings through the development of specific sensibilities and dispositions, leading people to orient and think about themselves within their environment in specific ways. A focus on how the projective flexibility of mind changes within different cognitive ecologies highlights interactivity, shifting material registers and affect. In this relational domain, we can begin to explore qualitative differences within digital tools and practices. We might analyse how digital technics fundamentally challenge notions of scale, knowledge and agency by revealing and making salient new aspects of the world around, but also how the digital might in turn hide as much by advancing sleek interfaces over deep knowledge and engagement. Potential gaps between emerging perspectives and human affective experience presents us with opportunities for developing notions of mind, materiality and embodiment for the study of digital design. Digital tools allow us to call up worlds of unprecedented complexity and layer and fold our own material and informational concepts and experiences into these. The long term and short term of cognitive evolution flow together through embodied material engagement with software and interfaces. Material engagement theory's commitment to a temporal and relational perspectives allows for functional analysis of how emerging technologies lead to qualitative changes in cognitive ecologies while situating these within larger societal and historic contexts. This relation between micro and macro historical perspectives might prove an exciting new venue for research in the history and theory of design and architecture.

For now, bringing together two instances of digital design in our analysis, we have shown how cognitive ecologies of practice have roots reaching well beyond the individual, while their evolution relies inextricably on imaginative engagement in the here and now. It should be obvious that intimate knowledge of ancient boatbuilding or the finer points of polynomial expressions is not a prerequisite for working with digital design software today. These were the exact things Bézier and others sought to overcome by establishing a more intuitive interface for digital design and manufacture. As material anchors and cultural and institutional practice has stabilised around these interfaces their manipulation has become a given, and entry-level CAD is taught on most architecture courses today. But, as mentioned in the introduction, more and more architects and educators are beginning to ask if rudimentary understanding of tools and interfaces is enough to steer the profession through a digital transformation (Oxman 2008). In light of the above it should be obvious that a lack of computational knowledge might hamper creative and imaginative work within the complex digital design ecologies of today. While the presets of contemporary design software offers powerful tools to shape and transform 3D creations, real originality comes from transcending presets and trying out new combinations when scripting or building models. This digital sense making relies on a deep understanding of computation and software—as well as morphology, materials and aesthetic experience. Burry (2011b, 2015) has been vocal about the danger he perceives in designers engaging uncritically with digital tools, fearing that this furthers a lack of understanding of the crafts and conceptual domains that make digital design possible in the first place. If this is indeed the case, digital tools might make design practice the poorer for it, as uncritical engagement with software impoverishes the rich cognitive ecologies that support creative design research and imaginative play at the human-computer interface (cf. Suchman 2006). There is nothing in digital tools per se that leads to this outcome; their use and designs are contingent on their coupling within wider cognitive ecologies. The power of digital tools lies not only in the way their memory, speed and precision allows us to control complex and dynamic systems, but also in their ability to conflate and combine a vast array of heterogeneous logics and information at the scale of human perception and experience. Digital design research requires some understanding of all of these realms. Embodied material engagement is where these come together to create meaningful connections through the ongoing sense making between designer, tools and materials. It might be futile to try to completely untangle these strains or isolate one or the other, but looking at their entwining through material engagement gives us an analytical tool for seeing how this shapes creative opportunities for the digital designer.

References

Archer, B. (1995). The nature of research. *Co-design, interdisciplinary journal of design*, 6–9.

Bézier, P. (1993). P. Bézier: How a simple system was born. In G. E. Farin (Ed.), *Curves and surfaces for computer aided geometric design: A practical guide* (pp. 1–12). Boston: Academic.

Bézier, P. (1986). A personal view of progress in computer aided design. *ACM SIGGRAPH Computer Graphics SIGGRAPH, 20*(3), 154–159.

Bézier, P. (1998). A view of the CAD/CAM development period. *IEEE Annals of the History of Computing, 20*(2), 37–40.

Bézier, P. (1983). UNISURF, from Styling to tool-shop. *Computers in Industry, 4*(2), 115–126.

Burry, M. (2015). Antoni Gaudí and his role in forming a traditional craft and digital culture continuum. *International Conference on Culture and Computing (Culture Computing), 2015,* 3–9.

Burry, M. (1993). *Expiatory church of the sagrada familia: Antoni Gaudi*. London: Phaidon.

Burry, M. (2011a). Geometry working beyond effect. *Architectural Design, 81*(4), 80–89.

Burry, M. (2005). Homo faber. *Architectural Design, 75*(4), 30–37.

Burry, M. (2011b). *Scripting cultures: Architectural design and programming*. Chichester, UK: Wiley.

Cross, N. (1982). Designerly ways of knowing. *Design Studies, 3*(4), 221–227.

Cross, N. (2006). *Designerly ways of knowing*. London: Springer.

Davis, D. (2013). Modelled on software engineering: Flexible parametric models in the practice of architecture (PhD Dissertation, Melbourne: School of Architecture and Design, RMIT University).

Farin, G. E. (2002). A History of curves and surfaces in CADG. In G. E. Farin, J. Hoschek, & M. Kim (Eds.), *Handbook of computer aided geometric design* (pp. 1–22). Amsterdam: Elsevier.

Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science, 22*(2), 133–187.

Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic.

Fauconnier, G. (1997). *Mappings in thought and language*. Cambridge, U.K.: Cambridge UP.

Frayling, C. (1994). Research in art and design. *Royal College of Art Research Papers, 1,* 1.

Gallagher, S. (2014). Pragmatic interventions into enactive and extended conceptions of cognition. *Philosophical Issues, 24,* 110–126.

Gosden, C., & Malafouris, L. (2015). Process archaeology (P-Arch). *World Archaeology, 47*(5), 701–717.

Hutchins, Edwin. (2005). Material anchors for conceptual blends. *Journal of Pragmatics, 37*(10), 1555–1577.

Hutchins, Edwin. (2010). Cognitive ecology. *Topics in Cognitive Science, 2*(4), 705–715.

Hutto, D. (2015). Overly enactive imagination? radically re-imagining imagining. *The Southern Journal of Philosophy, 53,* 68–89.

Hutto, D., Kirchhoff, M. D., & Myin, E. (2014). Extensive enactivism: Why keep it all in? *Frontiers in Human Neuroscience, 8,* 1–11.

Hutto, D., & Myin, E. (2013). *Radicalizing enactivism. Basic minds without content*. Cambridge MA: MIT Press.

Johnston, S. A. (1994). Matthew Baker and the art of the shipwright (PhD Thesis, Cambridge). *Making Mathematical Practice: Gentlemen, Practitioners and Artisans in Elizabethan* England, 107–165.

Kirsh, D. (2009). Projection, problem space and anchoring. In N. A. Taatgen & H. V. Rijn (Eds.), *31st annual meeting of the cognitive science society 2009 (CogSci 2009): Amsterdam, Netherlands, 29 July–1 August 2009* (pp. 2310–2315). Austin, TX: Cognitive Science Society.

Kirsh, D. (2013). Thinking with external representations. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain, computation, interactivity and human artifice* (pp. 171–194). London: Springer.

Lakoff, G., & Núñez, R. (2000). *Where mathematics come from: How the embodied mind brings mathematics into being*. New York: Basic.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UP: Cambridge.

Malafouris, L. (2011). Enactive discovery: The aesthetic of material engagement. In R. Manzotti (Ed.), *Situated aesthetics: Art beyond the skin* (pp. 123–141). Imprint Academic: Exeter.

Malafouris, L. (2013a). *How things shape the mind: A theory of material engagement*. Cambridge, MA: MIT.

Malafouris, L. (2013b). On thinking and form-making. In M. A. Brayer & F. Migayrou (Eds.), *Naturaliser architecture/naturalizing architecture* (pp. 244–253). Editions HYX: Collection FRAC centre.

Malafouris, L. (2014). Creative thinking: The feeling of and for clay. *Pragmatics & Cognition, 22* (1), 140–158.

Malafouris, L. (2015). Metaplasticity and the primacy of material engagement. *Time & Mind, 8*(4), 351–371.

Menges, A., & Ahlquist, S. (2011). *Computational design thinking*. Chichester, UK: Wiley.

Núñez, R. (2010). Enacting infinity: Bringing transfinite cardinals into being. In J. Stewart, O. Gapenne, & E. Di Paolo (Eds.), *Enaction: Toward a new paradigm for cognitive science* (pp. 307–334). Cambridge, MA: MIT.

O'regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(05), 939–973.

Oxman, R. (2008). Digital architecture as a challenge for design pedagogy: theory, knowledge, models and medium. *Design Studies, 29,* 99–120.

Oxman, R., & Oxman, R. (2013). Theories of the digital in architecture. Routledge.

Piegl, L. A., & Tiller, W. (1997). *The NURBS book*. Berlin: Springer.

Rogers, D. (2000) Introduction to NURBS: With historical perspective. Morgan Kaufmann.

Senseney, J. R. (2011). *The art of building in the classical world: Vision, craftsmanship, and linear perspective in greek and roman architecture*. New York: Cambridge UP.

Simon, H. A. (1996). *The sciences of the artificial*. Cambridge, MA: MIT.

Steffensen, S. V. (2013). Human interactivity: Problem-solving, solution-probing and verbal patterns in the wild. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain, computation, interactivity and human artifice* (pp. 195–222). London: Springer.

Suchman, L. (2006). *Human-machine reconfigurations. Plans and situated actions* (2nd ed.). Cambridge: Cambridge UP.

Terzidis, K. (2006). *Algorithmic architecture*. Oxford: Architectural.

Turnbull, D. (1993). The Ad Hoc collective work of building gothic cathedrals with templates, string, and geometry. *Science, Technology & Human Values, 18*(3), 315–340.

Vallée-Tourangeu, F., & Villejoubert, G. (2013). Naturalising problem-solving. In S. J. Cowley & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain, computation, interactivity and human artifice* (pp. 241–254). London: Springer.

Chapter 15
Organisational Cognition: A Critical Look at the Theories in Use

Davide Secchi and Billy Adamsen

Abstract This chapter is concerned with exploring the ontology of organisational cognition (OC) through conceptual mapping in order to recognise and understand what OC really is about. The objective is not to provide a comprehensive literature review of this area, but to map the concept so that both meaning and extent of its reach can be better defined. In so doing, the article considers several perspectives under which the domain of "organisation" interacts with or relate to "cognition" (or it does not do so). A table that summarises similarities and differences among approaches is presented. Finally, the table is then used as a tool to demonstrate overlaps, gaps, and define possible directions for future research in the OC field.

Introduction

Organisational cognition (OC) is an expression that has now been around for a few decades (e.g., Ilgen et al. 1994; Hodgkinson and Healey 2008). The aim of this chapter is to explore the meanings ascribed to these two words. In pursuing this aim, it becomes possible to provide a framework to categorise the existing theoretical approaches and compare them to the state-of-the-art in cognition research. This conceptual exercise should also lend itself towards suggesting ways to move the conversation forward.

D. Secchi (✉)
COMAC Research Cluster, Centre for Human Interactivity (CHI),
Department of Language and Communication, University of Southern Denmark,
Slagelse, Denmark
e-mail: secchi@sdu.dk

B. Adamsen
Zealand Institute of Business & Technology, Slagelse, Denmark

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8_15

Some of the inspiration for this chapter came from the first edition of this book (Cowley and Vallée-Tourangeau 2013). As we see it, this second edition is a way to map progress on the previous publication and, at the same time, indicate viable directions for cognition research. We do believe there is scope to extend cognition research to social environments and organisations that clearly represent an essential part of when and where cognition occurs (Secchi and Neumann 2016). Moreover, distributed and *systemic* views of cognition emphasise the interconnectedness of all parts of the cognitive process between internal, external, macro, meso, micro levels and why it is important to focus on *social* aspects (Hutchins 1995; Vallée-Tourangeau and Cowley 2013; Secchi and Cowley 2016; Neumann and Cowley 2016). The 'social' is one of three pillars that constitute the basis for distributed cognition research, as originally indicated by Hollan et al. (2000). Hence the idea of an enquiry on whether existing OC approaches incorporate some of them as well as brings the discourse forward in the direction of developing social cognition. For readers of this book, it is probably unnecessary to present a detailed description of systemic/distributed cognition here because other authors in this volume provide better and more focused analyses. However, given the broad range of topics that fall under this approach to cognition, it is probably useful to indicate the few that interest us in this chapter. Above all, research in OC has tended to be split between those who do and do not separate internal and external cognitive resources. While the 'internalist' focus is often linked to Simon's (1979) concept of bounded rationality, is in increasingly recognised that even this requires the 'smart interplay' of what is deemed internal and external (Clark and Chalmers 1998; Clark 2003). Accordingly, we refer to the complex dynamic systems or niches that emerge as a by-product of cognition 'in the wild' (Hutchins 1995). These systems are socially anchored (or distributed) and partially transcend single individuals, making cognition happen the way it does (Hutchins 2013). Elsewhere, we called these occurrences 'social organising' (Secchi and Cowley 2016).

The chapter takes a conceptual and a critical stance. We consider existing studies on OC and classify them into 'approaches' on the basis of a rather abstract conceptualisation of how the field of management deals with cognitive science. In this chapter we deliberately look for critical aspects in which management scholars have studied and continue to study cognition. The critical part of the chapter is clearly driven by the fact that we are examining the approaches in an attempt to see whether there is trace of a more current view of cognition, i.e. the systemic/distributed view mentioned above, of the type described in this and the previous edition of *Cognition beyond the brain*.

A Quick Look at Journal Publications

The idea to explore the cognitive side of organisational behaviour was introduced during the so-called "cognitive revolution" in management and organisational studies (MOS) in the early Nineties (Ilgen et al. 1994). Back then, the number of MOS scholars interested in cognition and conducting research that applied some

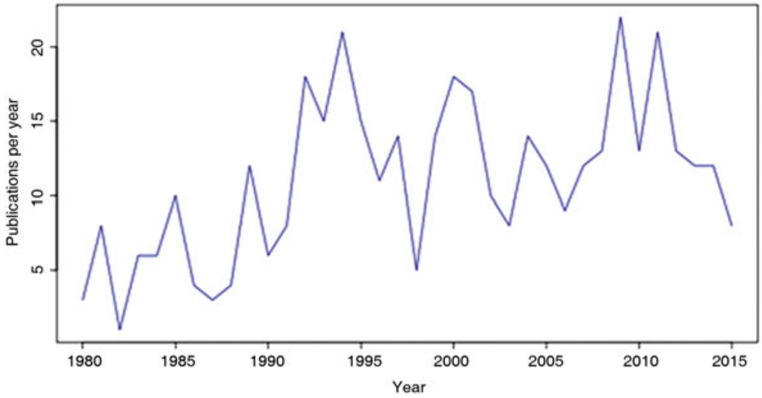


Fig. 15.1 Number of papers published per year on cognition-related topics in top-tier MOS (EBSCO, Business Source Complete database; 1980–2015)

aspects of cognition to organisations was on the rise. A quick review of articles published on the subject confirms this trend. Data in Figs. 15.1 and 15.2 are collected counting how many articles were published in any given year using 'cognition' as a subject term in EBSCO's Business Source Complete database. Articles published before 1980 were too dispersed over the years and for this reason not included. To understand how well OC research has been influential, we included most of the traditional MOS mainstream journals,[1] with particular reference to those dealing with decision making and aspects of psychology applied to the workplace (Fig. 15.1). We then compared these numbers to the overall number of articles appearing in the same business database and having 'cognition' as a subject term. Journals focusing mostly on cognition and psychology with no reference to business or organisations were excluded because they publish on the topic with a focus

---

[1] Here is our list: Academy of Management Review, Academy of Management Journal, Journal of Management, Journal of Management Studies, Organization Science, Organizational Behavior and Human Decision Processes, Journal of Organizational Behavior, Administrative Science Quarterly, MIS Quarterly, Journal of Applied Psychology, and Personnel Psychology. A mainstream journal is one that usually covers research topics that are commonly accepted as being part of a field or discipline. Its articles attract, on average, more citations than others and can be identified by bibliometric indices (e.g., Davis and Eisemon 1989; Nagpaul and Sharma 1994—both in Scientometrics). In the context of this research, a mainstream journal is one that ranks atop the latest Journal Citation Report by ISI Thomson for the discipline 'management'. We excluded journals unlikely to cover cognition because their aim seemed far from the topic (e.g., Journal of Information Technology, Omega) and those with too limited issues per year (e.g., Academy of Management Annals). We compared the 2014 ranking to previous years and the journals meeting these excluding criteria are also those that do not appear permanently in top positions. We also tried to include journals that traditionally cover cognition and have a high impact factor although they are not in the top ten (e.g., Organizational Behavior and Human Decision Processes, Journal of Organizational Behavior).

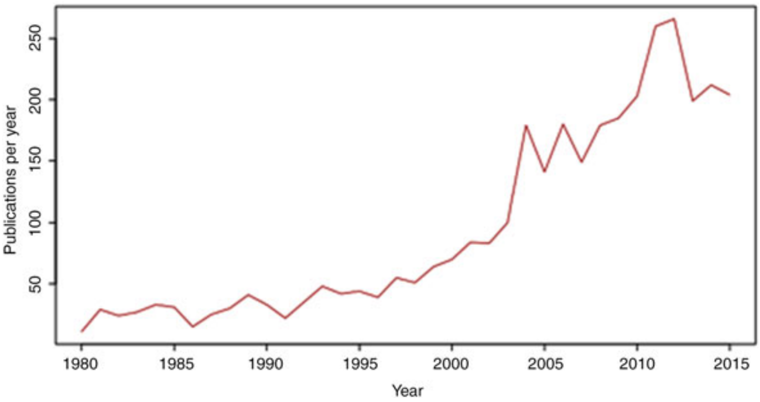


Fig. 15.2 Number of papers published per year on cognition-related topics in other business journals (EBSCO, Business Source Complete database; 1980–2015)

that is not related to MOS (Fig. 15.2). Yet, it is also important for us to emphasise that the data from these journals cannot be considered as an indication of a trend but rather as *waves* of publications.

There are at least three *waves* in the data, one in the early Nineties, another around year 2000, and the third around year 2010. The first seems to confirm the interest in the topic and this is in line with Ilgen et al. (1994) and also Walsh (1995). While this is almost exclusively a MOS phenomenon, the increase in journal publications in the third wave (i.e., around 2010) seems in line with an overall increase in cognition-related research in wider MOS and in other business fields. The second wave seems to be related to MOS research as well, although the overall number of articles published on the topic elsewhere was also increasing. However, more recently it appears that there has been a decrease in cognition-related publications, both in mainstream and other journals. It is probably too early to be definitive about this development as it may well be that it is a temporary decrease, just like it happened in the past decades. The number of publications in mainstream journals signals that the topic has entered the core of a scientific discipline and signals that cognition-related work is accepted as one of the legitimate areas of study in that discipline. As shown in this review, over the last 36 years an average of 15% of the papers on cognition in business-related journals was published in the MOS mainstream. Especially the year 1994 seems particularly important as it represents a turning point in the field. In that year ca. 33% of the papers were published in the top-tier journals selected for this study. This is a true peak that has not been repeated since, probably because of the exceptional number of publications and because it was still a pioneering topic back then. It can be claimed that 1994 was the year where the message that cognition was an acceptable topic to consider among others in the discipline was sent to the MOS community. If we split the data by decade, a slightly different picture emerges. The overall proportion of articles on OC published in this selection of mainstream journals is 17% for the

Eighties, followed by the Nineties with 23%. A slight decline can then be observed in the following decade with 10%, and conclude with the final 6 years (2010–2015), where the percentage is down to 5%. These numbers make sense only in relation to the average total number of articles because the actual number of publications on cognition-related topics in top-tier journals per year has been stable at 13 over the last three decades.

These numbers are particularly helpful when it comes to understand whether there is a case (or not) to be made for the study of cognition in organisations. From the numbers above it is reasonable to claim that indeed this is the case, and we can also show that there is a rather stable tendency to host cognition-related content in top journals as well as in more generally business-related outlets. However, there seems to have been little progress over the decades. In fact, the average number of articles published in mainstream MOS journals has been stable at 13 per year over the last three decades. This occurs despite the fact that the overall interest in cognition has been increasing in other business-related publications and in the field overall. Figure 15.2 shows that there has been a growth in the number of articles published every year, with the highest peak registered in 2012. Despite the fact that mainstream MOS journals tend to dedicate approximately 1.4% of their space to cognition every year, it is not in line with the general increase in the business literature that can be observed. However, what is undeniable is that there is a literature on OC in the MOS field which unquestionably justifies the scope of this chapter. Moreover, this also moves the enquiry to another related aspect of this subject/field: what is actually meant by the terms ‘organisational cognition’?

Organisational Cognition for Management Researchers

Shared interests brought some researchers affiliated with the US Academy of Management (AOM) to create an interest group—called a ‘division’—dedicated to the study of Managerial and Organizational Cognition (MOC). Their domain statement can be used as a preliminary attempt to understand where the interest lies within scholars working on MOC-related topics. From the website, we gather that they research “how organization members model reality and how such models interact with behaviors. Major topics include: attention, attribution, decision making, ideology, information processing, learning, memory, mental representations and images, perceptual and interpretive processes, social construction, and symbols” (AOM/MOC web page, 2016). The statement is formulated by academics affiliated with the division who work and publish within this area of study. This domain statement covers a significantly vast range of topics and it identifies the most traditional areas in which cognition has been applied—not only within organisations. While an interesting point to start from, in order to have a deeper appreciation of the boundaries of the domain, we turn to two seminal articles that served as review of the literature during the past two decades (Walsh 1995; Hodgkinson and Healey 2008). Both approaches take a psychological approach to cognition as it applies to organisations.

For Walsh (1995), the study of cognition in organisations is fundamentally dedicated to the representation, development, and use of so-called ‘knowledge structures’ that are analysed at the individual, group, and organisational level. To use Walsh’s own words, a “knowledge structure is a mental template that individuals impose on an information environment to give it form and meaning” (1995, p. 281). A second more recent literature review by Hodgkinson and Healey (2008) still labels the domain as ‘cognition in organisations’ and defines it through the approaches used to analyse it. Under this angle, cognition changes depending on one of the five of the following approaches: “(a) schema theory and related conceptions of mental representations (especially the notion of mental models), (b) behavioral decision theory (especially work on heuristics and biases), (c) attribution theory, (d) social identity theory and related conceptions, and (e) enactment and the related notion of sensemaking” (2008, p. 391).

Although the authors above do not provide definitions and these are really our understanding of the frameworks the two articles are based upon, we can still consider them as a starting point for this work. However, we can also raise a few concerns with some of the points that seem to shape the boundaries of this area of study. One concern is that, in defining the field, definitions inevitably set the boundaries of what can and cannot be considered under the label ‘managerial and organisational cognition’. This is a problem of most, if not all, (field) descriptions because they are created to set a limit for the field that, unfortunately, often finds its way into the very meaning of the term and the theory. In other words, when it comes to topics, descriptions are mostly exclusive and not inclusive because they set the bounds of what is a legitimate topic for a given field, area to focus on. In our case here, the limits seem to be particularly strict in that they seem to overemphasise the psychology of cognition while overlooking or underplaying other aspects. In particular and especially in Walsh (1995), there is an emphasis on a representationist view of cognitive processes. The explicit call for mental models and to the fact that cognition happens through internal representation of external knowledge is indicative of this tendency. As some pointed out recently (Chemero 2009), this is one way to approach cognition and certainly not the only viable way (more on this below). This may also lead to a clear divide between cognitive processing and behaviour. In this respect, ‘representationism’ becomes more restrictive, and less ‘distributed’, and significantly delimit the domain. Therefore, it places it in danger of sounding more like an ideological stance rather than an open field of enquiry.

The other concern is connected to the first because the description seems to be projected and modelled on well-established areas in the study of cognition. There is nothing particularly wrong with leaning on certainty and describing the field using the *status quo* except that it may signal that some opportunities are lost and some semantic elements in the term ‘organisational cognition' overlooked. Referring to opportunities and semantic elements we mean that ‘cognition in organisations’ can be intended as something different than what it has been in most traditional studies and to semantic elements that could nuance the definition of the term. In fact, putting cognition into organisations means to depart significantly from what studied in the last century. As indicated elsewhere (Secchi 2011), highly socialised environments—such as organisations—become pivotal in cognitive processing. These social aspects should be the core of OC as a field of study and a semantic element in the meaning of the term 'organisational cognition'. Consistently with this statement, we can then also claim that the means of studying cognition in organisations should be radically different from those used in the past, mostly considering individual decision making in isolation (as in, for example, Simon 1979; Kahneman 2003). The idea of cognition that this area of study seems to adopt is one that was developed within cognitive science, hence the tradition of artificial intelligence and lab-oriented research. Science gained tremendously from those fields although trans-disciplinary research is supposed to adjust, modify, and change from the 'home' field depending on what is relevant in the 'host' field. Do we have the means for a paradigmatic change in cognition research as it applies to organisations? While OC was introduced to management scholars in the Nineties, the field of cognition underwent another "revolution". In 1995 the book *Cognition in the Wild* (Hutchins 1995) appeared and, a few years later, the volume *Being There* (Clark 1998) was also published opening the field of cognition to *distributed* processes. This paradigm shift was mostly done to take into account the fact that the external environment plays a far more pervasive role in individual cognition than it was thought before (see other chapters in this book and in Cowley and Vallée-Tourangeau 2013). Moreover, the internal elements of cognition (i.e. the brain) is very much intertwined with external resources (Clark and Chalmers 1998). This chapter embraces a distributed and systemic approach to cognition and confront it to the work that MOC scholars have been publishing. This paradigmatic perspective allows us to circumstantiate the critical approach to current studies. At the same time, by the end of the chapter we will be able to understand whether MOC publications as represented by the various definitions above is consistent with more current views of cognition or not.

This chapter is not a literature review but uses a selection of articles published over the years on OC, assess their contribution and attempts to outline a possible direction in which this area of study can move forward. The work is structured as follows. We first define the criteria we used to identify the different types of contributions in this area, we then move from one class of study to the other and outline aspects that may (and/or seem to) hold cognition research back. This is not done for the sake of criticising the valuable research efforts produced in this area so far, but in the spirit of constructive criticism for the advancement of the discipline. In other words, this is done in an attempt to move from a simple gap spotting exercise in the current research paradigm to outline directions in which we can move forward (Alvesson and Sandberg 2011). Finally, we present a summary of the approaches that are covered, and set the ground for a more positive account of OC.

A Conceptual Map for Organisational Cognition

In an attempt to capture most ways in which OC has been studied, this chapter explores the area widely, considering how the two disciplinary domains—i.e., MOS on the one hand, and cognitive science, on the other hand—are or are not connected to each other in the selected studies. This taxonomy refers to possible theoretical frames/approaches that emerge when one looks at the two domain. The four frames/approaches selected for this article are actually represented in the literature, but our focus here is on their theoretical plausibility, advantages, and disadvantages. Instead of deriving the frames from a literature review we asked the question ‘how can one consider the two domains of cognition and organisation research’. The idea to explore the theoretical underpinnings of concepts and refer to some literature afterwards is used in conceptual articles (e.g., Burnes 2005; Secchi 2007) and it offers a few advantages. One is that it allows authors and, we hope, readers, to think broadly on a topic, exploring theoretical possibilities that are not limited to the existing literature. Another is that it forces to consider theoretical differences on a more abstract ground, making it easier to compare between them. Also, a frame/approach is not evaluated on a ‘bandwagon’ effect that depends on numbers of authors working in one direction, but more on its value-added—i.e. what does this frame/approach add to the discourse and, ultimately, to the field? A less studied approach may be more sound and consistent than a mainstream approach, judging from its theoretical grounds. We believe this makes our take more unbiased than more traditional literature review studies. Finally, the strength of this theoretical take on organisation and cognition is exactly on these two selected domains. By choosing to discuss the way organisation research and cognitive science are (or are not) linked together, we aim at exploring *how* the (non-)integration unfolds in practice. Our approach indicates many ways in which the (non-)relation is practiced theoretically and it is mirrored in the literature.

Using Venn diagrams, Fig. 15.3 shows the two domains represented as circles that overlap or stand alone, and the approaches to OC can be viewed as selective interest on either overlaps or autonomy/isolation. Although these approaches will be discussed in the remaining part of the chapter in relation to the literature, it is probably useful to describe them succinctly here:

(a) *The additive approach*. The first diagram in Fig. 15.3 (letter a) represents an approach that considers the two disciplines to be separate but additive. This means that the two maintain distinct methods, disciplines, and have limited points of contact because they do not formally talk to each other—this is the perspective of the so-called ‘rational’ systems theories in management (Scott 2003) in which cognition is studied as part of individual decision making (e.g., Simon 1979).

(b) *The combination approach*. The two entities are still separate but they share a common ground as well as keep separate domains (Fig. 15.3, letter b). This includes socio-psychological approaches to management (e.g., Weick 1995) and cognition as information processing.

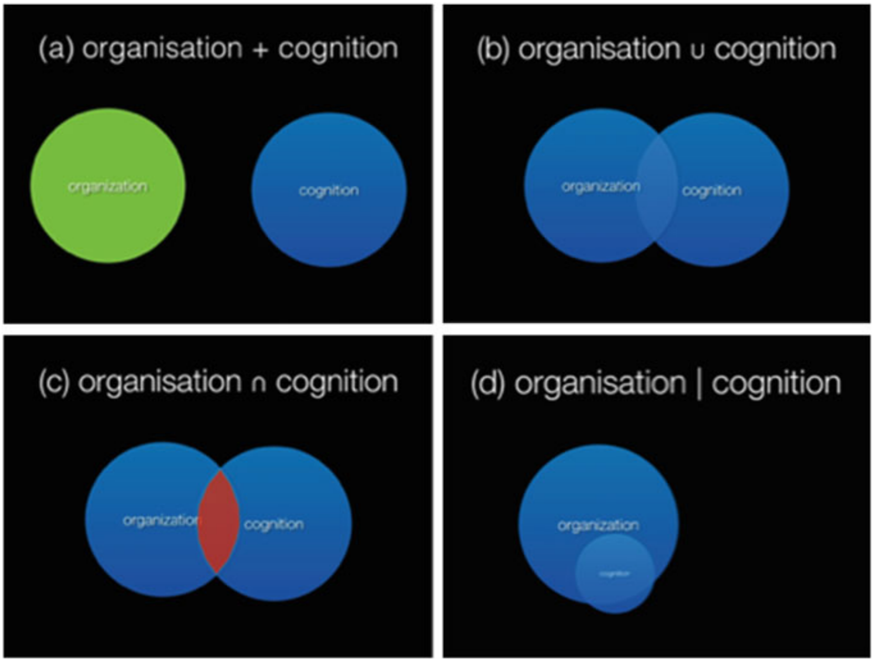


Fig. 15.3 A conceptual map of organisational cognition approaches

(c) *The intersection approach*. The intersection between the two domains is the only relevant element (Fig. 15.3, letter c), to indicate that something unique is defined when these two are together. The focus is on what people in organisations "share" (e.g., Cannon-Bowers and Salas 2001).

(d) *The conditional approach*. The most popular categorisation of OC probably is the one that considers cognition as one of the many elements that can be studied when conducting research in and of organisations (Fig. 15.3, letter d). This is what Hodgkinson and Healey cover in their review of the literature (2008).

This is a simplification of various positions and may not capture the approaches in greater detail. Also, what is represented in Fig. 15.3 does not necessarily reflect how scholars working in OC would conceptualise the framing of their views. This representation of theories/perspectives/approaches is, however, more a metaphor, in the sense that it is an abstract model that helps us highlight and discuss some salient aspects of the domain.[2]

---

[2] In this chapter the attention is limited to these four approaches, but the various combinations among the two diagrams in Fig. 15.3 can also include: (e)—*organisation*, an approach that uses the organisation as a conscious or unconscious benchmark to define cognition; (f)—*organisation U organisation*, this approach considers everything that is not organisation to define cognition but it then includes it back in; (g)—*cognition*, OC can be defined by everything but cognition and this is a way to define organisations as not affected by anything that can be related to cognition.

In the following we consider each one of these approaches, highlight their characteristics, and discuss how they develop and advance the study of cognition in organisations. As a guide to this analysis, we define a set of criteria that guide the enquiry. Each approach has an understanding of what 'organisation' is and how it reads through cognition or, vice versa, how there may be an understanding of cognition through a conceptualisation of what organisation is. Alternative to this cross-boundary perspective, the two domains may remain independent from each other. Stated differently, there may be instances where independence of the two domains is the characterising aspect of the approach, and some others where cross-fertilisation is the most relevant aspect instead. This is the first element considered through the analysis and is labelled as the *degree of domain dependency*. Another element is whether the approach is more cognition- or organisation-centred and this helps us define the degree of centrality that a more *socialised view* of cognition is taken into consideration. The last criterion relates to whether cognition is conceptualised as a simplistic or a more sophisticated set of mechanisms. We call this a *measure of complexity* for cognition research in organisations. The next section makes these aspects more explicit while the pages below deal with these criteria as a guide to the analysis.

An Additive Approach: Organisation + Cognition

In the *additive approach*, the two areas are kept separate. They can be added to each other although each one keeps distinct methods, disciplines, and share very limited points of contact. Hence, this view of OC is based on a lack of clear, immediate, and direct connection between the two domains. A clear definition of what OC is and what it aims at are beyond the objectives of most of those who conduct research using this approach. Moreover, a definition of the domain is marginal to these two areas of studies and is found only incidentally.

This approach is fundamentally not concerned with OC. Each discipline investigates phenomena within their respective domain, with little or no search for cross-fertilisation. This is the perspective of scholars working on organisation-related or cognitive-related domains who do not see the value of looking at their domain from, respectively, the cognitive or organisational viewpoint. Hence, the organisation remains a latent aspect of the cognitive activity analysed or, vice versa, cognition is a latent aspect of research focusing on organisational phenomena.

This is the case, for example, of the many studies that apply bounded rationality (Simon 1955, 1979) to show the limits of human reasoning (e.g., Kahneman and Tversky 1979; Gigerenzer and Selten 2001). Except for the more traditional approaches to bounded rationality (mostly from Herbert Simon), these studies are less concerned with organisational dynamics than the psychological conditions under which human rationality shows its limits. For this reason, the emphasis has been on *heuristics* and *biases*, mostly studied from a psycho-cognitive perspective. It was not intention of the authors that of understanding and/or defining these psycho-cognitive mechanisms specifically from an organisational viewpoint (although they sometimes refer to the 'social'; e.g. Kahneman 2003). Instead, they all looked at regularities in the way individuals behave and think independent of the given social context in which they operate. Yet, these two approaches have been often used to describe organisational phenomena (e.g., Bazerman 1994; Neale and Bazerman 1991); a few examples follow below.

In an article on innovation and organisational decline, for example, McKinley et al. (2014) use prospect theory to argue for an increase of innovation for managers when the firm is in decline. The need for something to 'save' the company from failure, bankruptcy, or death increases the likelihood of innovative thinking. When facing potential or actual losses, managers and employees would seek risks to confront the situation, and this is conducive of a greater degree of innovation (McKinley et al. 2014, p. 94).

In a foundational article on 'behavioural strategy', Powell et al. (2011) define the domain as merging "cognitive and social psychology with strategic management theory and practice" (p. 1371). The attempt starts with applications of prospect theory to business strategy, which emphasises cognitive biases and risk, among other reductionist takes (Powell et al. 2011).

Another take on cognition and prospect theory comes from the study of decisions under uncertainty compared to returns on investments (e.g., Kliger and Tsur 2011). The definition of a reference point is essential for managers to anchor their perceptions of losses and gains. Although the line of reasoning for this (and other) articles seem compelling, there seems to be a mismatch between the micro who/where/how of decision making and the macro measure for return that was selected from accounting statements. The article is valuable in that it highlights potentials for applications of prospect theory and it calls for further research.

Finally, there are other authors investigating how heuristics affect decisions in organisations. This is the approach taken, for example, by Artinger et al. (2015) who considers five classes of fast and frugal heuristics—decision processes when time and information are limited (e.g., Gigerenzer and Goldstein 1996)—and apply them to management. Authors show what are the practical and theoretical advantages of applying these tools to managerial decision making.

These studies share several assumptions and we mention two of the most relevant for the purpose of this chapter. First, authors operate on the grounds that individual decision making as well as individual cognition as it is studied in the lab could almost automatically be generalised to contexts other than those. Sometimes it does, sometimes it requires adjustments and caveats to be specified. As Powell et al. (2011) mention in their paper, these are reductionist approaches in the sense that they (a) are very traditional 'positive' perspectives within the social sciences and (b) tend to dissect the area of study to isolate meaningful effects. The implication is that a higher degree of adaptation is required if we are to transfer these studies to more applied and practical domains.

A second concern is that it is almost unanimously assumed that the social context is not particularly relevant although some mention explicitly that they are using a cognitive and social psychology perspective (Kahneman 2003; Powell et al. 2011). While the 'cognitive' and 'psychological' side of the research is extremely apparent, the social is almost completely missing.[3] This is of extreme concern because one of the most consistent ways to describe organisations has been through their 'social' attributes (Scott 1995, 2003). There are many outcomes that have potential to spin off such approaches, including the fact that researchers may be able to isolate specific cognitive mechanisms that apply to the 'average' man and woman and, in virtue of this, also to those who cover managerial positions. Yet, there is a clear *social* gap that somehow undermines the full transition of the approach to the study of organisations.

In summary, this approach does not conceive OC as a clear domain. Authors do see contributions of cognition in several areas of study but do not necessarily identify a new area of study with it; rather, the tendency is to claim that the original area is modified when adding cognition to it. This is the case, already mentioned above, of 'behavioural strategy' (Powell et al. 2011) where the adjective attached to the field identifies the fact that scholars bring in cognitive and psychological frames to study the domain.

Combination Approach: Organisation U Cognition

In this approach, the two areas are still separate but they establish contact and share some common ground. Although there is some cross-fertilisation, domains, methods, and subject are kept separate. What seems to be the case here is that the organisational domain takes what is needed from the cognitive domain.

The interesting aspect of this approach is that both disciplines borrow from each other although OC is not directly defined. However, concepts are just used instrumentally from the other domain and with the sole purpose of improving the domain of origin.

Amongst these approaches to organisation studies is Simon's *Administrative Behavior* (1997), where connections to decision making are done putting the seed for what was to come in later scientific productions. Along the same lines, the so-called information processing approaches (described in the initial literature review by Walsh 1995) use cognition-related concepts to analyse interactions among workers. This approach seems to interpret cognition as information processing, suggesting a very narrow understanding of it, typical of those who interpret

---

[3] We refer to the circumstance that sees the work of these scholars as a study of how a not better specified 'information' is used by the decision maker. Under this angle, it makes not difference whether the information is coming from another human being, from a computer, or any other social media. Instead, we claim this is exactly what makes organisations interesting. From a cognitive (and psychological) perspective, information coming from a social source makes the world of difference. Although partially inspired by Kahneman, Tversky and others, studies on advice giving and taking have started to unveils some of these aspects, at least from a psychological perspective (for a review, see Bonaccio and Dalal 2006).

it as mere computation, symbol manipulation, code crunching, and the like (Chomsky 1980; Fodor 1987; Newell and Simon 1972). Nevertheless, as one of the fathers of the field, Herbert A. Simon set the foundations of organisational behaviour on decision making and grounded it on cognitive processing (Simon 1997, 1979). His idea of bounded rationality (Simon 1955) is nothing but a model of individual cognition that came out by comparing existing economic theories with how managers actually make decisions in organisations (March and Simon 1958; Klaes and Sent 2005). The 'behavioural theory of the firm' is a micro-foundation of organisational phenomena where decision making becomes the key to all processes. Despite the combination of the two fields, Simon and the Carnegie school (Augier 2004; Gavetti et al. 2007, 2012) gained relatively limited terrain in explaining aspects of organising that are emergent, complex, dynamic and difficult to predict. This is not to be intended in a dismissive tone, quite the opposite we claim that this approach was successful in clearing the field from unrealistic assumptions (Gavetti et al. 2012). But it was not able to establish a true synergy between the two domains. This is probably more the legacy of March rather than that of Simon. In fact, in his career March has been interested in change—e.g., learning, ambiguity, and anarchy (March 1978, 1981, 1991; Cohen et al. 1972)—as a key feature of individuals in organisations. Even though he represents the progressive side of the school, the focus remains on the individual as if there is a conceptual bound that prevents the discourse to be extended farther. What we mean is that actual organisational phenomena were used to create more realistic assumptions for cognition and behaviour of economic actors, but these assumptions were not, then, projected back to the organisation. The 'social' is still sum of individual thinking and behaviour, where complexity is almost only hierarchical or structural (Simon 1962, 1976) rather than emergent, dynamical or social (Conte 1999). The combination between the two fields is there, their synergy is not.

In summary, this approach is consistent with the legacy of the early Herbert Simon (1997) and with those who hypothesised the behavioural theory of the firm (Cyert and March 1963). In an attempt to integrate the social back in, Simon introduced a model based on the disposition to lean on recommendations, advice, information to make decisions (Simon 1993). This attempt has had only limited attention in the literature so far but it has potentials to connect to more distributed views of cognition (Secchi 2011; Secchi and Bardone 2009).

The Intersection Approach: Organization $\cap$ *Cognition*

When the perspective of both organisation and cognition get closer, they define an overlapping area that exists because of the interaction between the two. For this reason, this conceptual area is labelled the *intersection* approach.

If the only relevant element here is the intersection between the two domains, it is then possible to use the convergence to indicate that something unique is defined. On the one hand, there are certain aspects of organisational life that can be only (or better) explained using cognitive lenses. On the other hand, there are some aspects of cognition that are inherently social and can be only (or better) explained using organisational lenses.

One of the areas of study that features within this particular intersection aspect is the tradition of "shared" cognition (Cannon-Bowers and Salas 2001). This is a concept that has been applied mostly to analyse teams and explain how team members share knowledge and information among each other. Shared cognition is not a uniform concept, it is more used to indicate a cloud of similar concepts that can all be framed in relation to the intersection domain—e.g., team mental models, transactive memory, etc. (Klimoski and Mohammed 1994). Nevertheless, there are common features that can be discussed here.

The shared cognition ground is intended for most as "knowledge structures held by members of a team that enable them to form accurate explanations and expectations for the task, and in turn, to coordinate their actions and adapt their behavior to demands of the task and other team members" (Cannon-Bowers et al. 1993, p. 221). From this definition it is clear that the interest lies in shared 'knowledge' by team members. This can be related to the task at hand and to the other team members' attitudes and beliefs (Cannon-Bowers and Salas 2001). It is also apparent that the conceptual framework is that of mental models (Johnson-Laird 1983). This implies that cognition is intended as a representation or a model of something that is outside the brain. Secondly, knowledge seems to be characterised as 'information' and anchored to the paradigm of cognition as information processing. Although this has its advantages in that it simplifies the treatment of cognition, it mostly focuses on the coding, decoding, and transfer of knowledge as information. This is apparent from the 'how' of shared cognition. In fact, Cannon-Bowers and Salas (2001) specify that shared knowledge should be compatible, identical, overlapping, and distributed (pp. 198–199). From such conceptualisation, we understand that the idea of knowledge can be identified objectively in order for it to be shared. If this is the perspective taken on cognition, even though the concept refers to teams, it is clear that the individual is still the core of the analysis. Knowledge is just the vehicle through which individuals exchange information and complete the tasks.

Shared cognition brought significant advancements to the study of how organisations complete tasks and achieve goals through teams (Healey et al. 2015). It also helped in understanding the dynamics of sharing knowledge among team members and in identifying the strategies to improve that exchange. Given the emphasis on knowledge as information processing, research focuses mostly on explicit, deliberate, and conscious activities. Recently, scholars have started to address the need for unconscious and implicit processes in shared cognition (Healey et al. 2015). By bringing in the implicit side of team member interaction there is room for the psychological element to be taken more seriously. However, this does not seem to be enough to overcome both a representationist (mental model) view and a reduction of cognition to information processing (Healey et al. 2015, p. 401). In summary, the study of shared cognition is the natural evolution of the 'rational' approaches to decision making in organisations (Simon 1997; Cyert and March 1963). Contrary to the other approaches reviewed above, this one has the ambition to define OC more directly by looking into teams. However, it fails to enrich the theory with more socialised aspects of cognition that are not retrievable simply adding parts to the system. In other words, we claim that the system (i.e., team cognition) should be more than the simple sum of interactive parts (i.e., team members). We challenge the view that a mix or merge of individual processing abilities provides a clear understanding of how teams deal with tasks. Among the many others, the shared cognition perspective falls short of explaining immaterial constructs such as cultural products, behavioural rules, routine formation, development, misuse of resources, role-bound behaviours, power, and how material artefacts affect thinking. Even if we consider only one of these aspects to be not marginal to the understanding of OC, we can appreciate the limits of shared cognition. We claim that these limitations are directly connected to the representationist and the information processing views of cognition. The former forces the cognitive discourse to be bound to models of reality, hence raising issues of accuracy. A more pragmatic approach would instead focus on how cognition happens, hence building a model ex post rather than ex ante (Chemero 2009). The latter is particularly problematic because it practically equates the human brain to a computer and it fundamentally treats cognition as some sort of input-output process (Miłkowski 2012). The sole consideration—among many others—that ‘humans have a body’ excludes treatment of cognition as clear-cut information processing because of emotions, ever-changing biological conditions, perception issues, medium-based interaction, and more (Varela et al. 1991). The most obvious critique here is not that shared cognition uses these traditional assumptions to the study of teams. Instead, it is that this perspective makes use of these *sole* assumptions as if cognitive science has not moved forward from Simon’s view (1997, first published in 1947).

Another and slightly more advanced view from the *intersection approach* has a more “psychological” or socio-psychological labelling. We refer to the theory of *sensemaking* in organisations (Weick 1995) that borrows from psycho-cognitive mechanisms to define how workers understand and elaborate on their environments. Since this approach has been widely employed in several areas of management (Gioia 2006), reviewing its main features is not within the scope of this chapter. We can, instead, characterise its relations with OC to explore how much it moves the field forward. According to Weick (1993), sensemaking was originally a reaction to the widespread emphasis on decision making within the MOS field of research. For the purpose of this chapter, only a few aspects are briefly reviewed here. First, sensemaking comes into play when a given phenomenon is coded into a language that is understood by organisational actors (Gioia et al. 1994). This implies—and this is a second aspect—that *meaning* is attributed only when there is an *explicit* and shared understanding of a given situation, action, process, or else in the organisation (Weick et al. 2005). In fact, as Weick et al. state, “[w]hen we say that meanings materialize, we mean that sensemaking is, importantly, an issue of language, talk, and communication. Situations, organizations, and environments are talked into existence” (Weick et al. 2005, p. 409). This is chiefly important because, on the one hand, language is a key aspect of human interaction and, on the other, it carries ‘meaning’ so that individuals are capable of making sense of the organisational world around them. Third, individual interrelations lead to the identification of a 'collective mind' as a disposition to heed (Weick and Roberts 1993). While the collective mind has a role that is separate from the individuals, only individuals contribute to the making of the collective mind. A collective can be a group, a team, a department, an organisation (Weick and Roberts 1993).

The sensemaking perspective brings the OC discourse closer to more recent developments in cognition. In particular, it introduces a more socialised view of individuals in organisations and it characterises cognitive activities as situated and enacted. In fact, it is the action of speaking to each other that allows individuals to attribute meaning of given situations. Also, we can fairly state that individuals are embedded in organisational situations when meaning formation happens. More recently, Whiteman and Cooper (2011) advocated for a more ecological perspective on sensemaking, calling for a recognition of the prominent role of the natural/physical environment as a recurrent aspect of individual experience. However, this reference to the ecological view seems questionable because it overlooks the most significant constituents of the environment: other human beings. Probably, the authors overlooked this element because they deemed other aspects to be at the core of organisational dynamics. However, others are too fundamental to be missed out in any definition of any environment. To our eyes, this 'ecological' approach is in danger of not being cognitively relevant because it is not systemic enough. This means that space should be integrated with time and the diverse levels at which the human cognitive system operates.

The sensemaking perspective helps to push the discourse forward, although scholars do not seem to be particularly concerned with OC. This is not an issue per se, but it becomes such when dealing with 'situated', 'enacted', or 'ecological' aspects. The claim that sensemaking is always situated is particularly important because the situation carries 'meaning'. In this respect, the call for the importance of the natural/physical environment introduces an ecological angle to this field of studies and contributes to improve its explanatory power. And yet, from an overview of its theoretical infrastructure, we have the impression that all these concepts operate on a separate basis. This is exemplified by the use of words such as 'mind' and 'meaning': they are vague expressions for something that seems difficult to grasp. On the one hand, the collective 'mind' concept is used to categorise aspects other than the individual that still refer to the individual (Weick and Roberts 1993). This is particularly relevant because it makes it difficult to explain dynamics that are somehow independent of the individual. Put differently, the collective mind as a disposition to heed seems to be too anchored to explicit and observable aspects of human interaction. On the other hand, the word 'meaning' can help solve this situation and include in the mix latent or hidden meanings. However, if hidden or latent they are not included in the collective while they can be in the way individuals use language. Language is ambiguous and lacks of a uniform 'meaning'. For this and other reasons, some prefer to refer to content.

Sensemaking has now grown into a well developed branch of MOS. However, we are yet to see how it contributes more directly to re-define the boundaries of cognition in organisations. Under this respect, one of the problems can be that the idea of cognition under which the approach is based remains extremely vague. This section has attempted to try and define what cognition is for Weick and the others who share this perspective, and our reading is that cognition equals sensemaking. If this is the case, then the collective mind can be equated to organisational cognition. The issue seems to be that instead of dealing with the problems concerning an extended and distributed view of cognition, the solution has been a change of focus, i.e. from cognition to sensemaking. There are some benefits to this swop but some of the important issues mentioned above remain open (more on this in the next section).

The Conditional Approach: Cognition as a Given

Figure 15.3d represents the fourth approach with two diagrams that intersect where the focus extends from one domain to the other. In particular, we take an OC perspective so that the diagram can be read as a true attempt to integrate the two domains—i.e. MOS and cognitive sciences. This OC-centred view makes us also focus mostly on how organisational scholars used cognition in their research. For this reason, the label on top of Fig. 15.3d reads '*Organisation|Cognition*', where the sign '|' means "conditional" and it is used to indicate that the second element in the formula is given to the first. From this representation, it is apparent that the approach relates to explaining organisational phenomena and takes cognition as an element used to make sense of them. Under this perspective, cognition is one among many other elements that can be considered when studying organisations. Hence, OC is yet another aspect of the organisational literature that emphasises the procedural and mechanical aspects of thinking (and how they connect to behaviour).

This is probably the most common conceptualisation in management and organisation studies. A very thorough literature review of cognition in organisations by Hodgkinson and Healey (2008) shows that there are many distinct elements that have been related to cognition taking a MOS angle. The common trait seems to be cognition as another element to be included in the mix of variables that explain organisational behaviour—this one being the most prominent aspect of interest. Hence, this approach features the most traditional MOS concepts and its associates them with bits of cognitive processing theory.

Among the many, the most obvious areas covered by this perspective are probably leadership (Lord and Emrich 2001) and motivation (Latham and Pinder 2005). This is because both concepts have a very long tradition in MOS and draw on psychology, hence they can be connected easily to cognition. There are two points of view that can be used in both areas. One pertains to the individual that is exercising leadership or that is motivated to perform a certain task, job or else. The other relates to the impact that leadership exercises on others, the team, and/or the organisation as a whole; the same can be said for the study of motivation as it affects others, team members, and/or the company.

Under the first angle, cognition has been identified chiefly as intelligence—intended as the "speed and depth of information processing" (Mumford et al. 2015, p. 302).

This approach has a long tradition in leadership research and can be counted among trait theories, substantiating the claim that cognitive abilities are an essential characteristic of leaders (e.g., Hiller et al. 2011). With motivation, instead, cognition traditionally took more the form of rationality and has been used in the form of rational expectations models to explain how employees are encouraged by rewards, bonuses, and other monetary incentives (Locke and Latham 2002). This also implies that there is only extrinsic motivation (Ryan and Deci 2000) that is considered using this observation point.

The second angle allows researchers to see cognition as a tool to make leadership more effective in that it illustrates how business leaders dynamically engage with uncertainty and threats (Vessey et al. 2011). Through the lenses of cognition, motivation is also defined by its impact on team performance (DeChurch and Mesmer-Magnus 2010) and it becomes more tied to shared cognition and transactive memory (see above).

There are many examples of this use of cognition. From a more technical angle, this approach seems to be characterised by the fact that something defined as cognition is considered as a determinant of managerial, organisational, or work-related phenomena. Cognition usually works as an independent variable, sometimes as mediator or moderator (e.g., DeChurch and Mesmer-Magnus 2010), while managerial, organisational or other work-related phenomena are (taken separately) the dependent variable of the model (e.g., DeChurch and Mesmer-Magnus 2010). In this approach there is little elaboration on what the theoretical underpinnings of the relation are; the overall impression is that there is an emphasis on testing and on empirical validation of various models of cognition.

As far as limits are concerned, we can certainly mention that these studies intend intelligence, on average, as a very traditional logic-based and computation-like concept (e.g., Mumford et al. 2015). Within this same framework, there can be multiple 'intelligences' hence cognition need not be associated with just one of them. Among the many, emotional intelligence has been connected to leadership as one of the abilities essential to establishing and maintaining relationships with team members and employees (Lemmergaard and Muhr 2013). The interesting aspect of considering emotions is that it implies an embodied idea of cognition—i.e., the inseparable role of the body's biology and chemistry on brain activity. In other words, the emphasis on the most traditional among intelligences—i.e., the logical, rational, and computational—makes it easier to narrow down cognition into something that can be treated quantitatively. However, it also reveals that this rather strict and old-fashioned interpretation of cognition limits the potentials for the understanding and analysis of its applications to MOS research and practice. In particular, a computational view is strictly connected to the rationalistic foundations in organisations (see above; Simon 1997; Cyert and March 1963). What is the problem with a strict computational view? There are at least two major problems. One is that it considers psychological and social processes as ancillary at best; the other is that it reduces the 'thinking' to symbol manipulation, code cracking, and input-output mechanisms. Both problems combined make it extremely hard to tie cognition to dynamic and/or change processes leading to the emergent elements of organising such as cultures, rule making, and routines. Finally, from the MOS literature briefly reviewed above, we can see that cognition is considered a given to understand specific aspects of Organisational life. The fact it is framed as ‘given’ also points at the lack of discussion around what cognition is and what it is not, including the limits of using a narrow and old-fashioned view.

Comparing the Approaches

The interest in our classification becomes more useful when we compare the approaches over a set of criteria. As noted above, the first assessment criterion is the degree of *domain dependency*, the second is *centrality*, and the third criterion relates to the level of *complexity*.

Domain dependency is particularly important to understand whether the approach under consideration is benefiting of cross-fertilisation, or it is borrowing from the other without changing significantly. The perspective used in this classification is that of MOS and not that of cognition. In other words, we infer a measure of fertilisation that the MOS domain experiences through the use of cognitive science. After reviewing some of the publications falling into one of the approaches described above, we ask the question whether the field has changes its ways of understanding, analysing, or doing science due to cognitive science cross-fertilisation. The question can be answered with a three-point scale where each approach is assessed on a low, medium, or high level of cross-fertilisation.

The second aspect concerns *centrality* and it signals whether the approach focuses mostly on cognitive aspects of organising or it remains anchored to traditional organisational approaches. Put differently, we claim that there is an inner core to each approach that is either falling in the organisation or in the cognition domain. Depending on where the centre is located, there is a clearer understanding of the perspective used to frame cognitive and organisational phenomena.

Finally, the third aspect relates to the extent to which a simple or more complex idea of cognition is taken into consideration. This aspect is framed using five elements inspired by the so-called 4E cognition, an approach for which cognition is *embedded*, *enacted*, *embodied* (e.g., Menary 2010), and *ecological.*[4] We add a fifth dimension for classification purposes, and that is information processing (or computational) cognition, a dimension most approaches seem to lean upon. In summary, the claim is that a more complex view of cognition includes most, if not all,

---

[4] Menary (2010) leaves the ‘ecological’ out and substitutes it with the ‘extended’; although the extended is also very important, we believe it is rather a constituent allowing distributed system to work rather than one of the characteristics of its processes. Magnani (2007) for example, characterises two basic constituents as ‘externalising’ and ‘re-projecting’ to describe the ‘smart interplay’ (Clark and Chalmers 1998). As it is apparent in the text above, we are not discarding the extended features of cognition, rather emphasising those processes that would be particularly important to take into consideration with OC.

of these five aspects. While the information processing and computational view of cognition should be evident to most readers at this point of the chapter, together with the *embodied* view (mentioned above), the remaining three call for an explanation. The word *embedded* refers to cognition being un-detachable from resources external to the brain. This includes all those material and immaterial artefacts that human beings find themselves in, from buildings to language, from computers to culture. The assumption is that cognition cannot possibly work without the interchange between internal and external resources (e.g., Clark and Chalmers 1998). In a more inclusive view, embeddedness includes also the social domain and other human beings (Secchi 2011). Cognition is also *enacted* in the sense that it requires implementation and application from those involved in the process. This has been developed in different directions but one of the most interesting contribution comes from Magnani (2007) where he shows how reasoning about morality is inextricably linked to the way individuals behave. It is the action and the doing that shapes and integrates cognition in a way that is substantial to its existence. Finally, cognition is *ecological* because it creates ‘niches’ where individuals develop affordances with artefacts and other human beings. This leads to the development of emergent properties of the system in which cognition happens and explains, for example, culture (Hutchins 2013; Cowley 2016).

Now, why are we using 4E cognition to characterise complexity in cognitive thinking? The reason is obvious to us, but it may not necessarily be to our readers. By reviewing the organisational literature we came to the conclusion that most of the scholars working on OC are still anchored to a traditional view of cognition that places unnecessary limits on the field. In fact, the number of cognition-related articles published in MOS journals remained constant over the years (see Figs. 15.1 and 15.2), i.e. it has not increased. This signals, above all other aspects, that MOS scholars are doing the best they can with a concept that has almost unanimously been dismissed in its simplest form by most cognitive scientists and that is, at best, only one of many views of cognition. The information processing and computational paradigm is still dominating the field although it does so in relation to at least one of the elements in 4E cognition. In short, the choice of 4E cognition is made on the basis of (a) its clear indication of more current/contemporary elements of cognition research, and (b) the ability to connect to the more traditional computational approaches. These three criteria guide our understanding of the approaches and indicate a possible direction for further research in OC.

From Table 15.1 it is apparent that the most advanced approach among those reviewed is the one that considers the overlap between the two areas, i.e. the intersection approach. Moreover, within this approach, the most advanced model of cognition is the sensemaking perspective. All the others are anchored to a rather static and old-fashioned view of cognition where the social and the dynamics of interaction are overlooked. Overall, the picture is rather loom although consistent with our concerns regarding the definitions reviewed in the initial section of this chapter. This probably calls for a *new deal* in the cognitive revolution in management and organisation research.

Table 15.1 Comparing approaches to organisational cognition

Approach	Dependency	Centrality	Complexity				
			comp.	embe.	enact.	embo.	ecol.
Additive approach ($O + C$)	Low	C	✓				
Combination approach ($O \cup C$)	Medium	C	✓				
Intersection approach ($O \cap C$)	High	O	✓	✓	✓		
Conditional approach ($O \mid C$)	Medium	O	✓			✓	

Note O: organisational; C: cognition; comp.: computational; embe.: embedded; enact.: enacted; embo.: embodied; ecol.: ecological

Moving Forward

If there is ground for renovation in cognitive research applied to organisation research, this final section of the chapter tries to briefly outline how it can be carried over.

One of the points of this chapter so far has been that cognition in organisations as analysed above provides, at best, a partial perspective, and it often uses outdated cognition research. What we wrote so far opens the discussion to the understanding of the approach that should be taken when studying OC. We claim that this area of study would benefit from taking all the elements of *systemic/distributed* cognition (also labelled 4E for more practical purposes) into consideration at the same time. If we take this claim seriously then we should adopt a *radical systemic perspective*; it is radical because it takes a view that is very much distant from the current approaches to OC, and it is systemic because the emphasis is on complex social dynamics rather than on individuals or environments. By taking this perspective, we have the opportunity to define OC through social relations, dynamics and relationships. It is the interaction among these three aspects that, on the one hand, shapes organisational cognitive processes and, on the other, redefines what the organisation is about. Having a system view means exactly this, that is to say that priority is given to the dynamic aspects of social relations and relationships—hence power, roles, positions, responsibilityresponsibility, friendship, comradeship, identities, etc. This brings the focus on to emergent processes in organising, and these become apparent at many different scales within the organisation. This aspect leads, in turn, to two others. Firstly, the emphasis on emergence and complex adaptivity opens up for new methodologies to be used in the development of OC theory. As argued elsewhere (Neumann and Secchi 2016) agent-based computational simulations are particularly suited to support theoretical explorations in the social sciences and, in particular, in the MOS field. Secondly, this approach unveils one of the core contributions that OC would give to the social and cognitive sciences combined, i.e. a re-definition of relevant cognitive processes on social grounds. This was an intuition of Herbert A.

Simon who, towards the end of his career, acknowledged that there was the need to make decision making and cognition anchored to a more socialised (Simon 1993) or socially-based (Secchi 2011; Secchi and Bardone 2013) view of how individual think and behave. This view has already been introduced in the study of morality (e.g., Magnani 2007), chance seeking (Bardone 2011), diffusion of innovation (Secchi and Gullekson 2016), and learning (Miller and Lin 2010). In short, we believe that a radical systemic perspective provides the framework for OC studies to move forward rapidly and substantially.

Taking a different point of view, it can be seen that dynamics and relations recall the importance of both space and time. Some of the most significant organisational aspects that are usually overlooked by the OC approaches reviewed above are the long term effect of organising and their impact on individuals. We are referring to slow-timescale events (Neumann and Cowley 2016) such as culture, routine formation, and identity formation. Interesting and promising research could look at how these slow-timescale events (Neumann and Cowley 2016) impact, affect, modify and are modified by fast-timescale such as task completion, goal setting, motivation tools, position-based interactions, or roles, just to name a few. In between these two there is a plethora of meso-scale occurrences that is named *social organising* elsewhere (Secchi and Cowley 2016).

Figure 15.4 takes the diagrams above and presents a different way of drawing the circles together to indicate what a *radical systemic perspective* brings to the discourse. We still have the two circles, one for the ‘organisation’, the other for ‘cognition’. This time, though the two circles are almost entirely overlapping, meaning that cognition is everywhere in the organisation and that it becomes very difficult to disentangle it from organisational human activities. If we take that the two circles are one at the top and the one at the bottom, the eclipse is not full because there are domains in which the effect of one or the other area is blurred or not entirely relevant.[5] In the overlap, the two circles produce a few spin-offs, represented in Fig. 15.4 by other circles. These are to indicate that something ‘new’ or different is generated by the combination of these two circles, so that dynamics, relations, as well as time and space (Secchi and Cowley 2016) are considered as constituents of cognition in organisations or, we shall write now, *organisational cognition*.[6]

This study has limitations that help define its scope and reach. For example, the classification used requires a significant level of abstraction and is not necessarily the best way of conducting a critique. It is also artificial in that it forces studies to

---

[5] This is, for example, a computer-generated message or an automatic production process that follows a routine that once required human intervention but it is currently independent from it. That would characterise the process as mostly organisational and probably only indirectly related to cognition. Vice versa, there are purely cognitive processes that cannot be related to the organisation.

[6] A complete coverage of the model of organisational cognition through a radical systemic perspective is out of the scope of this chapter but it can be retrieved from a paper that was recently presented at the European Academy of Management conference by Secchi and Cowley (2016).

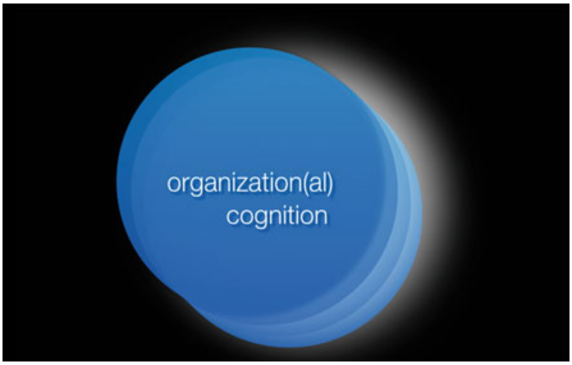


Fig. 15.4 A *radical systemic perspective* on organisational cognition

fall either within one or the other approach as they are described at the beginning of this chapter and in the pages above. However, some studies may fall in two categories. For example, some studies of shared cognition (letter c in Fig. 15.3) are also concerned with leadership and teams (letter d in Fig. 15.3). Another limitation relates to the fact that the choice of the four approaches may seem somehow arbitrary. It was guided by the number of studies that could be found on each given perspective. Note 2 specifies the other domains that we decided not to include in this chapter.

Conclusions

This chapter reviewed some approaches to OC by taking the perspective of organisational scholars. In so doing, we took a conceptual and a critical stance, starting from evidence available from scholars working on cognition in organisations. The conceptual stance allowed us to define four approaches, identified on the basis of the relation that organisation and cognition have. When the two stay separate, an *additive approach* is possible. Instead, when the two are in contact but maintain distinct domains, we have a *combination approach*, as opposed to an *intersection approach*, that is when scholars focus on the overlapping domain. Finally, the *conditional approach* identifies those studies where cognition is yet another variable used to explain organisational behaviour. We found that most of these approaches lean on traditional conceptualisations of cognition as information processing and none of them embraces a truly systemic and distributed perspective. This leaves an enormous ground for moving this field forward on the basis of a *radical systemic perspective* that would define OC on the basis of social relations, dynamics and relationships. This perspective seems promising not only for the advancement of the organisational field but for the cognitive sciences as well.

References

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review, 36*(2), 247–271.

Artinger, F., Petersen, M., Gigerenzer, G., & Weibler, J. (2015). Heuristics as adaptive decision strategies in management. *Journal of Organizational Behavior, 36,* S33–S52.

Augier, M. (2004). March'ing towards 'a behavioral theory of the firm' James G. March and the early evolution of behavioral organization theory. *Management Decision*, *42*(10), 1257–1268.

Bardone, E. (2011). *Seeking chances. From biased rationality to distributed cognition*, volume 13 of Cognitive Systems Monographs. New York: Springer.

Bazerman, M. H. (1994). *Judgement in managerial decision making* (3rd ed.). New York: Wiley.

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes, 101*(2), 127–151.

Burnes, B. (2005). Complexity theories and organizational change. *International Journal of Management Reviews, 7*(2), 73–90.

Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior, 22,* 195–202.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and Group Decision Making* (pp. 221–246). Hillsdale, NJ: Lawrence Erlbaum.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Clark, A. (1998). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Clark, A. (2003). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. Oxford, UK: Oxford University Press.

Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis, 58,* 7–19.

Cohen, M. D., March, J. G., & Olsen, H. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly, 17*(1), 1–25.

Conte, R. (1999). Social intelligence among autonomous agents. *Computational and Mathematical Organization Theory, 5*(3), 203–228.

Cowley, S. J. (2016). Cognition beyond the body: Using ABM to explore cultural ecosystems. In D. Secchi & M. Neumann (Eds.), *Agent-Based Simulation of Organizational Behavior. New Frontiers of Social Science Research* (pp. 43–60). New York: Springer.

Cyert, R. M. & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs: Prentice-Hall.

Davis, C. H., & Eisemon, T. O. (1989). Mainstream and non mainstream scientific literature in four peripheral Asian scientific communities. *Scientometrics, 15*(3–4), 215–239.

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology, 95*(1), 32–53.

Fodor, J. A. (1987). *Psychosemantics. The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.

Gavetti, G., Greve, H. R., Levinthal, D. A., & Ocasio, W. (2012). The behavioral theory of the rm: Assessment and prospects. *Academy of Management Annals, 6*(1), 1–40.

Gavetti, G., Levinthal, D., & Ocasio, W. (2007). Neocarnegie: The carnegie school's past, present, and reconstructing for the future. *Organization Science, 18*(3), 523–536.

Gigerenzer, G. & Goldstein, D. (1996). Mind as a computer: Birth of a metaphor. *Creativity Research Journal, 9*(2–3), 131–144.

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality. The adaptive toolbox*. Cambridge, MA: MIT Press.

Gioia, D. A. (2006). On Weick: An appreciation. *Organization Studies, 27*(11), 1709–1721.

Gioia, D. A., Thomas, J. B., Clark, S. M., & Chittipeddi, K. (1994). Symbolism and strategic change in academia: The dynamics of sensemaking and influence. *Organization Science, 5,* 363–383.

Healey, M. P., Vuori, T., & Hodgkinson, G. P. (2015). When teams agree while disagreeing: Reflexion and reflection in shared cognition. *Academy of Management Review, 40*(3), 399–422.

Hiller, N. J., DeChurch, L. A., Murase, T., & Doty, D. (2011). Searching for outcomes of leadership: A 25-year review. *Journal of Management, 37,* 1137–1177.

Hodgkinson, G. P., & Healey, M. P. (2008). Cognition in organizations. *Annual Review of Sociology, 59,* 387–417.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. In *Proceedings of the ACM Transactions on Computer-Human Interaction (TOCHI)* (Vol. 7 pp. 174–196). New York: ACM Digital Library.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (2013). The cultural ecosystem of human cognition. *Philosophical Psychology, 27*(1), 34–49.

Ilgen, D. R., Major, D. A., & Spencer, L. (1994). The cognitive revolution in organizational behavior. In J. Greenberg (Ed.), *Organizational behavior: The state of the science* (pp. 1–22). Mahwah, NJ: Lawrence Erlbaum.

Johnson-Laird, P.N. (1983). *Mental Models*. Cambridge, MA: Harvard University Press.

Kahneman, D. (2003). A perspective of judgement and choice. Mapping bounded rationality. *American Psychologist, 58*(9), 697–721.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Klaes, M., & Sent, E.-M. (2005). A conceptual history of the emergence of bounded rationality. *History of Political Economy, 37,* 27–59.

Kliger, D., & Tsur, I. (2011). Prospect theory and risk-seeking behavior by troubled firms. *Journal of Behavioral Finance, 12*(1), 29–40.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management, 20,* 403–437.

Latham, G. P., & Pinder, C. C. (2005). Work motivation theory and research at the dawn of the twenty-first century. *Annual Review of Psychology, 56,* 485–516.

Lemmergaard, J., & Muhr, S. L. (2013). *Critical perspectives on leadership: Emotion, toxicity, and dysfunction*. Cheltenham: Edward Elgar.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*(9), 705–717.

Lord, R. G., & Emrich, C. G. (2001). Thinking outside the box by looking inside the box: Extending the cognitive revolution in leadership research. *The Leadership Quarterly, 11*(4), 551–579.

Magnani, L. (2007). *Morality in a technological world. Knowledge as a duty*. New York: Cambridge University Press.

March, J. G. (1978). Bounded rationality, ambiguity and the engineering of choice. *Bell Journal of Economics, 9,* 587–608.

March, J. G. (1981). Footnotes on organizational change. *Administrative Science Quarterly, 26*(4), 563–577.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science, 2*(1), 71–87.

March, J. G., & Simon, H. A. (1958). *Organizations*. New York: Wiley.

McKinley, W., Latham, S., & Braun, M. (2014). Organizational decline and innovation: Turnarounds and downwards spirals. *Academy of Management Review, 39*(1), 88–110.

Menary, R. (2010). Introduction to the special issue on 4E cognition. *Phenomenological Cognitive Science, 9,* 459–463.

Milkowski, M. (2012). Limits of computational explanation of cognition. In V. C. Muller (Ed.), *Philosophy and theory of artificial intelligence* (pp. 69–84). Heidelberg: Springer.

Miller, K. D., & Lin, S.-J. (2010). Different truths in different worlds. *Organization Science, 21*(1), 97–114.

Mumford, M. D., Watts, L. L., & Partlow, P. J. (2015). Leader cognition: Approaches and findings. *The Leadership Quarterly, 26*(3), 301–306.

Nagpaul, P., & Sharma, L. (1994). Research output and transnational cooperation in physics subfields: A multidimensional analysis. *Scientometrics, 31*(1), 97–122.

Neale, M. A., & Bazerman, M. H. (1991). *Cognition and rationality in negotiation*. New York: Free Press.

Neumann, M., & Cowley, S. J. (2016). Modelling social agency using diachronic cognition: Learning from the Mafia. In D. Secchi & M. Neumann (eds.), *Agent-Based Simulation of Organisational Behavior. New Frontiers of Social Science Research* (pp. 289–310). New York: Springer.

Neumann, M., & Secchi, D. (2016). Exploring the new frontier: Computational studies of organizational behavior. In D. Secchi & M. Neumann (Eds.), *Agent-Based Simulation of Organizational Behavior. New Frontiers of Social Science Research* (pp. 1–16). New York: Springer.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Powell, T. C., Lovallo, D., & Fox, C. R. (2011). Behavioral strategy. *Strategic Management Journal, 32,* 1369–1386.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic denitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67.

Scott, W. R. (1995). Symbols and organizations: From Barnard to the institutionalists. In O. E. Williamson (Ed.), *Organization theory. From Chester Barnard to the present and beyond* (2nd ed.). New York: Oxford University Press.

Scott, W. R. (2003). *Organizations. Rational, natural, and open systems* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Secchi, D. (2007). Utilitarian, managerial, and relational theories of corporate social responsibility. *International Journal of Management Reviews, 9*(4), 347–373.

Secchi, D. (2011). *Extendable rationality. Understanding decision making in organizations*. New York: Springer.

Secchi, D., & Bardone, E. (2009). Super-docility in organizations. An evolutionary model. *International Journal of Organization Theory and Behavior, 12*(3), 339–379.

Secchi, D., & Bardone, E., (2013). Socially distributed cognition and intra-organisational bandwagons: Theoretical framework, model, and simulation. *International Journal of Organisation Theory and Behavior, 16*(4), 521–572.

Secchi, D., & Cowley, S. J. (2016). Organisational cognition: What it is and how it works. In *European Academy of Management Annual Conference*, *Paris, France*.

Secchi, D., & Gullekson, N. (2016). Individual and organizational conditions for the emergence and evolution of bandwagons. *Computational and Mathematical Organization Theory, 22*(1), 88–133.

Simon, H. A. (1955). A behavioral theory of rational choice. *Quarterly Journal of Economics, 69* (1), 99–118.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society, 106*(6), 467–482.

Simon, H. A. (1976). How complex are complex systems. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, volume Two (pp. 507–522).

Simon, H. A. (1979). Rational decision making in business organizations. *American Economic Review, 69*(4), 493–513.

Simon, H. A. (1993). Altruism and economics. *American Economic Review, 83*(2), 156–161.

Simon, H. A. (1997). *Administrative behavior* (4th ed.). New York: The Free Press.

Vallee-Tourangeau, F., & Cowley, S. J. (2013). Human thinking beyond the brain. In S. J. Cowley & F. Vallee-Tourangeau (Eds.), *Cognition beyond the brain. Computation, interactivity and human artifice*. London: Springer.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Vessey, W. B., Barrett, J., & Mumford, M. D. (2011). Leader cognition under threat: "just the facts". *The Leadership Quarterly, 22*(4), 710–728.

Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization Science, 6*(3), 280–321.

Weick, K. E. (1993). The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly, 38*(4), 628–652.

Weick, K. E. (1995). *Sensemaking in organizations* (1st ed.). Thousand Oaks, CA: Sage.

Weick, K. E., & Roberts, K. H. (1993). Collective mind and organizational reliability: The case of flight operations on an aircraft carrier deck. *Administrative Science Quarterly, 38,* 357–381.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science, 16*(4), 409–421.

Whiteman, G., & Cooper, W. H. (2011). Ecological sensemaking. *Academy of Management Journal, 54*(5), 889–911.

Index

S.J. Cowley and F. Vallée-Tourangeau (eds.), *Cognition Beyond the Brain*,
DOI 10.1007/978-3-319-49115-8